1-1-2016

Essays in the Economics of Health, Risk, and Behavior

Adam Leive
University of Pennsylvania, leive@wharton.upenn.edu

Essays in the Economics of Health, Risk, and Behavior

Abstract
The first chapter examines consumer choices of health insurance contracts. An important innovation in health insurance design is a high-deductible health plan paired with a health savings account (HSA). These contracts aim to control costs by linking insurance coverage with tax incentives for saving, but their rules are highly complex. How consumers perceive the features of these contracts may dampen any cost reduction and produce unintended welfare effects by distorting plan choices. Using a novel administrative dataset linking health insurance choices, medical claims, and saving in HSAs and 401(k)s from a large U.S. health insurer, I develop and estimate a model that integrates HSA saving with deductible choices. I estimate over two-thirds of the marginal HSA dollar is allocated to reduce the deductible, which counteracts the contract's cost-control incentives and leads consumers to choose different insurance plans than they would without an HSA. In this setting, using HSA contributions to offset higher deductibles produced no reduction in health care costs. Several counterfactual analyses quantify the welfare implications of using the HSA to finance current costs on moral hazard, plan enrollment and premiums, and the consumption smoothing benefits from insurance. Health insurance contracts that require sophisticated consumer decision-making may work well in theory, but may be less effective and lead to unintended consequences in practice.

The second chapter investigates how status affects health by comparing mortality between Gold and Silver medalists in Olympic Track and Field. Contrary to conventional wisdom, winners die over two years earlier than losers. Analysis of individual Census records of each athlete and his parents suggests that income is the key mechanism: losers pursued higher-paying occupations than winners after the Olympics, while parental earnings in childhood were similar. An athlete's performance relative to expectations plays an auxiliary role, but is much less important than income. The results suggest that how people respond to pivotal life events can produce long-lasting consequences for health.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Health Care Management & Economics

First Advisor
Jonathan Kolstad

Subject Categories
Economics

ESSAYS IN THE ECONOMICS OF HEALTH, RISK, AND BEHAVIOR

Adam A. Leive

A DISSERTATION

in

Health Care Management and Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

Jonathan Kolstad, Assistant Professor of Economic Analysis and Policy
University of California at Berkeley

Graduate Group Chairperson

Eric Bradlow, Professor of Marketing

Dissertation Committee

Iwan Barankay, Associate Professor of Management
Mark Pauly, Professor of Health Care Management
Robert Town, Professor of Health Care Management

ESSAYS IN THE ECONOMICS OF HEALTH, RISK, AND BEHAVIOR

ACKNOWLEDGEMENTS

I am grateful to my dissertation advisors Iwan Barankay, Jonathan Kolstad, Mark Pauly, and Robert Town for guidance and mentorship. Many people at Penn and elsewhere provided helpful comments and suggestions on different chapters of this dissertation. Guy David provided much encouragement during my studies and the job market process. I also appreciate the conversations and feedback I have received from a number of other faculty members in the department. I particularly wish to thank my classmates, from whom I learned tremendously during my studies. I am grateful to the department for enabling me to pursue my PhD at Wharton and financially supporting my studies. Additional funding from the National Institute of Aging through Grant Number T32-AG000186 to the National Bureau of Economic Research, from the Social Security Administration to the Center for Retirement Research at Boston College, and from the Ackoff Fellowship from the Wharton Risk Management and Decision Processes Center is gratefully acknowledged. Finally, I am eternally grateful to my family and friends for their constant support and encouragement.

ABSTRACT

ESSAYS IN THE ECONOMICS OF HEALTH, RISK, AND BEHAVIOR

Adam A. Leive

Jonathan Kolstad

The first chapter examines consumer choices of health insurance contracts. An important innovation in health insurance design is a high-deductible health plan paired with a health savings account (HSA). These contracts aim to control costs by linking insurance coverage with tax incentives for saving, but their rules are highly complex. How consumers perceive the features of these contracts may dampen any cost reduction and produce unintended welfare effects by distorting plan choices. Using a novel administrative dataset linking health insurance choices, medical claims, and saving in HSAs and 401(k)s from a large U.S. health insurer, I develop and estimate a model that integrates HSA saving with deductible choices. I estimate over two-thirds of the marginal HSA dollar is allocated to reduce the deductible, which counteracts the contract's cost-control incentives and leads consumers to choose different insurance plans than they would without an HSA. In this setting, using HSA contributions to offset higher deductibles produced no reduction in health care costs. Several counterfactual analyses quantify the welfare implications of using the HSA to finance current costs on moral hazard, plan enrollment and premiums, and the consumption smoothing benefits from insurance. Health insurance contracts that require sophisticated consumer decision-making may work well in theory, but may be less effective and lead to unintended consequences in practice.

The second chapter investigates how status affects health by comparing mortality between Gold and Silver medalists in Olympic Track and Field. Contrary to conventional wisdom, winners die over two years earlier than losers. Analysis of individual Census records of each athlete and his parents suggests that income is the key mechanism: losers pursued higher-paying occupations than winners after the Olympics, while parental earnings in childhood were similar. An athlete's performance relative to expectations plays an auxiliary role, but is much less important than income. The results suggest that how people respond to pivotal life events can produce long-lasting consequences for health.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Health Insurance Design Meets Tax Incentives: Consumer Responses to Complex Contracts

1.1. Introduction

Recent efforts to curb U.S. health care costs have centered on consumer decision-making in insurance. An important and innovative health insurance contract is a high-deductible health plan (HDHP) paired with a health savings account (HSA). Compared to traditional insurance plans, HDHPs feature higher deductibles to encourage consumers to be more cost-conscious in their health care use. In return for the higher deductible, consumers gain access to an HSA—a unique tax-preferred savings vehicle that functions like a 401(k) retirement plan, with the additional feature that withdrawals for health care expenses are always tax-exempt. Many policymakers and economists have championed these contracts as a market-based solution to efficiently reduce health care costs.[1] Collectively named “consumer-directed health care,” HDHPs and HSAs now represent one-quarter of enrollment in the employer health insurance market and are popular on many state insurance exchanges (Kaiser Family Foundation 2015; Herman 2014). The Affordable Care Act’s “Cadillac Tax” and limits on employee premiums will further accelerate HDHP adoption by discouraging more generous coverage.

However, the impact of HDHPs and HSAs on costs and welfare depends on the way consumers perceive the features of these contracts, which are more complex than other insurance products. The policy rationale of combining a high deductible with a savings account is to provide consumers with protection from catastrophic risks, while

[1] For example, see Pauly and Goodman (1995); Council of Economic Advisors (2004); Feldstein (2006); Bush (2006); Daniels (2010); Cogan, Hubbard, and Kessler (2011).

increasing their sensitivity toward moderate costs because their own money is at stake and all unused HSA funds roll over. This "use it or save it" provision of HSAs is a key distinguishing feature from the "use it or lose it" provision of Flexible Spending Accounts, but many consumers may be unaware of this attribute. In addition, HSAs are the only health insurance contract that permits using tax-preferred saving for non-health consumption, starting at age 65. In theory, this feature helps forward-looking consumers internalize the total cost of care because purchasing current health care with HSA funds subtracts from future income, bringing HSAs closer to the first-best contingent-claims contract that smooths consumption without causing moral hazard (Arrow 1963; Zeckhauser 1970; Feldstein 2005). In practice, though, consumers may heavily discount the future or fail to recognize the substitutability between health care and other goods from HSA saving. So while employers and insurers intend for the contract to reduce costs through a higher deductible, consumers may view HSA saving as a way to reduce the deductible, blunting the contract's cost-control incentives.[2]

In this paper, I examine insurance plan and saving decisions to test whether consumers use the HSA's tax incentives to reduce the current deductible or to save for future consumption. I extend a discrete choice model of insurance plans by integrating HSA saving with deductible choices. There are two key parameters in this model. The first is risk aversion, which is modeled as a random coefficient. The second is a "marginal propensity to consume" (MPC) parameter measuring the proportion of the marginal HSA dollar saved to finance current out-of-pocket costs versus future consumption. This parameter summarizes the potential effect of various economic

[2] A growing literature documents challenges in consumer-decision making in health care, including understanding plan features (Loewenstein et al. 2013; Handel and Kolstad 2015), the dynamics of prices when faced with non-linear contracts (Liebman and Zeckhauser 2008; Abaluck et al. 2015; Dalton et al. 2015; Aaron-Dine et al. 2015), plan choices in complex choice environments (Abaluck and Gruber 2011; Kling et al. 2012; Heiss et al. 2013; Bhargava et al. 2015), and the way prices are framed on menus (Schmitz and Ziebarth 2015).

fundamentals on insurance plan and utilization choices, including discounting, mental accounting, and information frictions. For example, if consumers use HSAs to reduce the deductible because of preferences for current consumption, this parameter will capture discounting. If consumers view HSA funds as reserved for health care rather than as fungible, this parameter will reflect mental accounting. While the parameter may capture various optimization errors, there may also be rational reasons why consumers fund HSAs to reduce short-term costs.[3] Regardless of the underlying economic mechanisms, what I call the MPC summarizes how HSAs influence insurance plan choice and health care consumption, similar to how the marginal propensity to consume out of income determines the fiscal multiplier in a macroeconomic model. This new model of insurance choice nests the standard model as a special case, when the MPC equals zero.

Identification of the choice model relies on an exogenous switch to HDHP coverage, variation in prices (e.g. premiums, deductibles, tax rates, matching rates), and information about preferences for retirement saving (e.g. 401(k) choices). Conditional on 401(k) saving, HSA saving should not differ across insurance plans if consumers view the HSA as a retirement savings vehicle. In that case, insurance choices would be equivalent to those from a standard model in which risk aversion is the only source of preference heterogeneity. By contrast, a positive relationship between HSA saving and the deductible chosen indicates consumers use the HSA partly to reduce current health care costs. Observing 401(k) saving allows me to estimate the model's parameters without estimating a fully structural life-cycle model of savings. I assume any unobservables—such as beliefs about income processes or bequest motives—and preferences for future consumption operate equivalently between the HSA and 401(k)

[3]For example, people may use the HSA to finance routine expenses or elective surgery—services they would have consumed without an HSA but are now tax-subsidized.

given the nearly identical structure and rules between the two accounts. Importantly, this approach does not require assumptions about the optimality of 401(k) saving, which may be violated in practice (Choi et al. 2011).[4]

I estimate this new model of insurance plan choice using a novel administrative dataset on health insurance decisions, HSA saving, and 401(k) saving from a large U.S. health insurer. This insurer replaced its traditional low-deductible insurance offerings exclusively with HDHPs for its own employees and provided a menu of HDHPs to choose between. In the years following the switch, the employer adjusted the menu of its insurance deductibles, premiums, and matching rates for HSA contributions, providing identifying variation in prices. I exploit variation in insurance premiums to identify risk aversion separately from the MPC and variation in 401(k) saving to identify the MPC separately from risk aversion. Detailed panel-level data includes insurance deductible choices, contributions by both employees and employer to the HSA and 401(k), medical and pharmacy claims of the employees and any dependents, demographics, and information on salary and job characteristics. The choice model incorporates distributions of health care costs at the time of plan and saving decisions, which are constructed using the claims and other employee-level observable characteristics. These detailed data, price variation, and the exogenous switch to HDHPs/HSAs provide for clean identification of the choice model, which I estimate via simulated maximum likelihood.[5]

[4]For the same reasons it is difficult to determine the optimality of 401(k) savings, I do not focus on the optimality of HSA saving. However, there are some HSA saving decisions that are unambiguously sub-optimal following revealed preference arguments. People who make 401(k) contributions in excess of the employer match should be contributing the limit to their HSA, since the tax benefits of the last HSA dollar are weakly better than for the last 401(k) dollar. In this setting, only 7.5 percent of employees in this situation max out their HSA, providing some evidence that many people make optimization errors.

[5]As described in detail in Section 4, premiums vary cross-sectionally by level of deductible (from a menu of four choices) and family composition, and vary longitudinally as the employer raised the deductibles offered.

I find that HSA saving is chosen primarily to reduce the current deductible, rather than to save for future consumption. Sixty-eight cents of the marginal HSA dollar is allocated towards paying for current out-of-pocket costs, with a 95 percent confidence interval between 65 and 72 cents. HSA saving thus is largely used to counteract the contract's high deductible. In this setting, replacing traditional coverage exclusively with HDHPs produced no reduction in health care costs because employees offset the higher deductibles with HSA contributions. The model estimates risk aversion to be 3.5×10^{-4} on average, in line with previous literature, and is positively correlated with the MPC. As a robustness test, a supplementary calibration using additional data on HSA withdrawals obtains similar parameter estimates as the choice model. The calibration takes advantage of the fact that HSA saving offers a continuous choice of insurance deductible, in effect, when HSA contributions are chosen to reduce the deductible. Taking the level of each observation's HSA withdrawals as a measure of contributions for current costs, the choice of deductible then point-identifies risk aversion for each household conditional on its cost distribution and marginal tax rate. Median risk aversion is estimated to be 2.4×10^{-4}, close in magnitude to the choice model estimates. The MPC is calculated as the ratio of HSA withdrawals to HSA assets, with 69 cents of every HSA dollar being withdrawn in a given year. This calibration provides supporting evidence that consumers treat HSAs primarily as a way to reduce their deductible.

Using the HSA to pay for current costs has important implications for contract choices, changes in health care consumption, and insurance premiums, which I illustrate through several counterfactual exercises. Each counterfactual treats the MPC as "welfare-relevant" in that the parameter affects consumer utility once enrolled in a contract. In HDHPs with an HSA, a positive MPC affects consumption decisions and plan choices by serving as a price reduction. I consider two different counterfactual plans as benchmarks: (1) a traditional insurance plan with 20 percent coinsurance, representing the status quo; and (2) an HDHP without an HSA (in which the MPC equals zero by construction), representing how the insurance deductible was intended to function.

Changes in the perceived price of the HDHP from HSA saving can be translated into moral hazard effects. If the demand curve reflects the marginal benefit of care, then the model's estimates imply that HDHPs with an HSA are only 28 percent as effective at reducing moral hazard compared to an HDHP alone. In particular, an HDHP with HSA reduces moral hazard by $125 compared to a $450 reduction from an HDHP alone. Using the HSA to pay for current costs thus substantially counteracts the cost control objective of the high deductible.[6] At the same time, the welfare impact of HSAs depends on consumer understanding of the health benefits of care. If consumers misjudge the benefits from care or make other optimization errors in their health care consumption decisions—which Baicker et al. (2015) define as "behavioral hazard"—then the demand curve is no longer sufficient to quantify moral hazard. Consumers without knowledge of the health benefits from care may actually be worse off if they perceive the full price of care under the HDHP since they may reduce services worth more than their cost. In the most comprehensive study to date on health care consumption patterns with HDHPs, Brot-Goldberg et al. (2015) document consumers cut all services, including valuable preventive care, once enrolled in HDHPs. Depending on the extent to which consumers make such mistakes, HSA saving may lead to health improvements relative to an HDHP without an HSA that offset increases in moral hazard. In addition, reducing the deductible through HSA

[6]There is an established literature on the welfare consequences of moral hazard and tax subsidies for insurance premiums beginning with Arrow (1963), Pauly (1968) and Feldstein (1973) and reviewed by Finkelstein (2014).

contributions provides \$387 in consumption smoothing benefits, on average, compared to an HDHP alone, but such gains are modest compared to the mean tax expenditure of \$591 on HSA contributions.

For a given menu of HDHP options, perceiving the HSA as a price reduction on the deductible also leads to different HDHP choices. I demonstrate the influence of HSA saving on deductible choices through a counterfactual that predicts market shares of an HDHP with and without an HSA and the average costs of consumers selecting each plan. In this setting, HSA saving attenuates any adverse selection based on health risk compared to an HDHP without an HSA. Specifically, the incremental costs of the set of consumers choosing more generous coverage relative to the highest deductible is lower when contracts include an HSA. These differences translate into premiums for more generous coverage often being \$1,000 lower for contracts with an HSA than ones without an HSA.

This paper contributes to several strands of literature. First, it adds to a growing literature on health insurance choices.[7] By incorporating saving decisions into insurance choices, the paper extends the standard insurance choice model beyond claim probabilities and risk aversion, similar to recent work in health (Handel 2013; Handel and Kolstad 2015) and other contexts (Sydnor 2010; Barseghyan et al. 2013). My model demonstrates the importance of incorporating HSA saving with deductible choices when estimating consumer risk preferences and predicting plan choices. In this setting, almost 30 percent of consumers switch insurance plans over a two-year period.[8] A standard model in which risk preferences, which are time-invariant, con-

[7]See e.g. Cardon and Hendel 2001; Carlin and Town 2009; Einav et al. 2010; Abaluck and Gruber 2011.

[8]This rate is high considering no plan was dominated, in contrast to other research documenting lower switching rates in settings with dominated insurance plans (Handel 2013; Bhargava et al. 2015).

stitute the only source of unobserved preference heterogeneity struggles to account for such switching between plans. Instead, the standard model rationalizes choices by estimating risk aversion that is implausibly large in magnitude, on average, and with large plan- and time-specific shocks. Accounting for HSA contributions helps predict choices, leading to a lower estimate of risk aversion and a substantial improvement in the model's fit to the data. The paper also relates to recent work documenting challenges in consumer decision-making over non-linear health insurance contracts (Liebman and Zeckhauser 2008; Abaluck et al. 2015; Dalton et al. 2015) and when faced with complex choice environments (Kling et al. 2012; Bhargava et al. 2015). The study adds to existing research on HDHPs that have focused on consumption responses, rather than HSA saving, and documented moderate cost reductions (Buntin et al. 2011; Borah et al. 2011; Haviland et al. 2015; Brot-Goldberg et al. 2015). This study shows that how consumers use HSAs to respond to higher deductibles can fully offset any spending decline. In this way, the paper particularly complements Brot-Goldberg et al. (2015) who find consumers react to the deductible in other ways that undermine the contract's objectives.

1.2. Background on HSAs

Created with the 2003 Medicare Modernization Act, an HSA is a portable financial account that must be paired with a HDHP. In 2016, the statutory minimum deductible—the amount paid by the consumer before coverage begins—for a HDHP was $1,300 for self coverage and $2,600 for family coverage. HSA account holders cannot use their HSA if they switch to other types of health insurance while working. HSA contributions roll over each year, unlike Flexible Spending Accounts (FSAs) where the enrollee loses his unused balance at year's end.[9,10] Medicare beneficiaries can make HSA withdrawals, but not contributions.

Health savings accounts, which are owned by individuals, offer a powerful savings vehicle to finance both health care expenses and consumption in retirement. In many ways, HSAs closely resemble 401(k)s as shown in Table 1.1: contributions to HSAs are deductible from taxable income ("above the line"), interest grows tax-deferred, and withdrawals for non-medical consumption are subject to income taxation and a penalty if before age 65.[11] HSAs also offer survivor benefits, similar to 401(k)s.[12] However, HSAs provide superior tax advantages to 401(k)s since withdrawals at any age to finance qualified medical care are not counted as taxable income. Qualified expenses, which must be incurred after the HSA has been established, includes out-of-pocket payments associated with a HDHP while working, as well as premiums for Medicare, COBRA, or long-term care insurance.[13] HSA balances cannot be used tax-free to finance premiums for employer health insurance or Medigap (supplemental insurance for Medicare's out-of-pocket payments). Another tax advantage relative to 401(k)s is that employee contributions made as payroll deductions are not subject

[9] In October 2013, the Treasury Department announced FSA balances up to $500 could be rolled over from one year to the next. My study studies pre-2010 choices, before this rule went into effect.

[10] FSAs are compatible with traditional insurance and do not need to be paired with a HDHP as HSAs do. HSAs are also distinguished from Health Reimbursement Arrangements (HRAs) and Archer Medical Savings Accounts (MSAs). HRAs are owned by the employer, not the employee, and cannot be invested. The employer funds the account for qualified expenses and may decide whether HRA funds roll over from year to year or are forfeited at year's end. HRAs do not have to be paired with a HDHP. Archer MSAs, created as a pilot program in the mid-1990s, are more similar to HSAs in that funds roll over, accounts are portable, and HDHP insurance is required. Eligibility for Archer MSAs was restricted to self-employed or small employers (fewer than 50 employees) and the pilot program ended in 2007.

[11] The penalty for early withdrawals was 10 percent before 2011, equal to that of 401(k) withdrawals, and increased to 20 percent beginning in 2011.

[12] HSAs obtained through employers may or may not qualify as an ERISA plan, depending on the employer's involvement in the plan. HSAs not opened through an employer are considered personal savings vehicles by the Department of Labor and not protected under ERISA.

[13] Until 2011, over-the-counter drugs without a prescription were included as qualified medical expenses. Starting January 1, 2011, a prescription was needed for over-the-counter drugs to be financed tax-free with HSA funds.

to FICA taxes (Social Security and Medicare) and employer contributions are not subject to FICA or FUTA (unemployment insurance) taxes. [14]

In 2015, the annual HSA contribution limit including both employer and employee contributions was $3,350 for self-only coverage and $6,650 for family coverage. These limits, which the IRS increases over time for cost of living adjustments, rose from $2,850 and $5,650, respectively, in 2007. Individuals over age 55 can also make annual "catch-up" contributions of an extra $1,000 in 2015 to their HSAs, up from $800 per year in 2007.[15] By way of comparison, the IRS's annual limit on employee 401(k) elective deferrals was $18,000 in 2015 and employers could provide an additional $53,000 in defined contributions to an employee's 401(k) in the same year.

While HSA saving limits are substantially lower than 401(k) limits, the costs of health care in retirement is large enough for the average consumer to consider HSAs as sensible tool for long-term saving. On average, Medicare beneficiaries spend close to $5,000 annually out-of-pocket on premiums, long-term care facilities, and other services (Cubanksi et al. 2014). The net present value of out-of-pocket expenses not covered by Medicare at age 65 has been estimated at between $220,000 and $376,000 dependent on the time period and whether long-term care is included (Fronstin et al. 2008; Webb and Zhivan 2010; Yamamoto 2013; Fidelity 2014). Since there is a distribution around these costs, risk-averse workers will want to save more than these

[14] A policyholder can also pay out-of-pocket (with taxable savings or current income) for health expenses while enrolled in an HSA, save the receipts, and reimburse herself from the HSA at any time in the future, even decades later. In effect, this strategy transforms part of the HSA balance into a Roth IRA because taxes are paid on the out-of-pocket payments while working, interest earned on HSA balances is tax-exempt, and the HSA withdrawal in the future is tax-exempt.

[15] Between 2004 and 2006, annual HSA contributions were limited to the lesser of the statutory maximum or the chosen deductible. This restriction biases contribution and deductible choices to be positively correlated. This rule was repealed on December 20, 2006 by the Tax Relief and Health Care Act of 2006. As described in Section 4, I exclude employees who opened an HSA prior to 2008 from my analysis to avoid introducing possible correlation between deductible and saving choices induced by legislation.

averages to guard against high cost realizations. For example, Webb et al. (2010) estimate that out-of-pocket costs exceed $570,000 with a five percent probability in 2009—the last year of my sample period. In order to finance these costs with HSA funds, a consumer would need to fully fund his HSA without any withdrawals while working. This comparison is important because it indicates the average consumer can meaningfully trade-off current versus future health care consumption in making HSA choices.[16]

Since their creation in 2004, HSAs have grown to include 14 million accounts and assets of $25 billion, with the majority of accounts opening since 2011 (Fronstin 2014). The take-up of HDHPs and HSAs has increased dramatically over the last decade, covering 20 percent of people who obtain insurance through their employer (KFF 2014). Over one-quarter of firms now offer an HDHP/HSA option, and nearly half of large firms (with over 1,000 workers) do. Despite their increased popularity, contributions, have been modest. Based on the Employee Benefits Research Institute (EBRI) database comprising one fifth of accounts and assets nationwide, the average contribution (including employer contributions) in 2014 was about $2,000 for self and family coverage combined.[17] Roughly 10 percent of account holders contributed the maximum amount (Fronstin 2015a). On average, account balances amounted to $2,077 in 2014, up from $1,320 in 2007. Among accounts open at least five years, the average balance was $3,092 (Fronstin 2015a). The large majority of accounts is not invested in financial markets (Fronstin 2015b).[18] In a first study using tax records,

[16]Although there is some debate about whether Americans are prepared for retirement (see e.g. Engen et al. (1999), Scholz et al. (2006), Hurst (2008), Munnell et al. (2014))), there is consensus that risks to health represent an important component in retirement saving decisions. Both the direct cost of poor health in terms of medical care and the indirect costs, such as the inability to substitute home production for purchased goods, reduce assets of retirees (Skinner 2007; Poterba et al. 2010; DeNardi et al. 2010).

[17]Half of HSA accounts receive employer contributions (Fronstin 2014).

[18]Keeping HSA funds in low-yield saving accounts may reflect account rules on minimum balances or a lack of consumer information. However, since money is fungible, such behavior may actually be

Helmchen et al. (2015) find older and higher-income workers opened and fully funded their HSAs more often than did younger and lower-income workers.[19]

Evidence of how HDHPs affect health care costs is still emerging, but it points to moderate declines in spending compared to traditional insurance of around 15 percent (Buntin et al. 2011; Borah et al. 2011; Haviland et al. 2015; Brot-Goldberg et al. 2015). In one setting, the move from no out-of-pocket payments to an HDHP produced spending cuts across all services (Brot-Goldberg et al. 2015). Many consumers do not appear to be forward-looking even over the course of a year: those with low shadow prices based on their expected spending relative to the deductible reduced spending while below the deductible, thus reacting to spot prices instead. If such demand responses reflect behavioral hazard as well as moral hazard (Baicker et al. 2015), then such documented consumption drops cannot quantitatively estimate a reduction in moral hazard.

Limited research exists on how employees fund HSAs in relation to 401(k)s or how HSAs interact with deductible choices. Parente and Feldman (2008) find a weak positive correlation between contributions to HSAs and other tax-deferred retirement saving vehicles among one set of University employees. Yet, their sample included just 63 HSA accounts from a sample of 16,000—a take-up rate of just 0.4 percent—and their results were not robust to alternative specifications. Analyzing over 160,000 accounts held at United Health Group and OptumHealth Bank, Chen et al. (2013) also took a descriptive approach and found HSA contributions were negatively correlated with employer contributions and positively correlated with age, income, education,

a financially savvy investment strategy if consumers can rebalance their 401(k) and IRA by selling bonds and cash, keeping their total asset allocation unchanged, because HSAs provide option value before retirement in terms of tax-free withdrawals for medical care.

[19] Many states exclude HSA assets in determining Medicaid eligibility, so low contributions among lower-income workers are likely not due to consideration about Medicaid enrollment.

and health care spending. They estimated a positive but insignificant correlation with the deductible, and they lacked data on 401(k)s or other saving accounts. Peter and Steinorth (2012) simulated a life-cycle model where health care spending and lifespan were stochastic (with certain labor income), but made the assumption that individuals max out contributions, which is at odds with observed contribution levels. Other studies on HSAs focused on the choice of insurance plan, comparing traditional insurance to a high-deductible health plan (Cardon and Showalter, 2007; Steinorth, 2011; Handel, 2013; Handel and Kolstad, 2015), but did not study saving decisions. Cardon and Showalter (2001) discuss the role of FSA contributions as supplemental insurance coverage, which also applies to HSAs, and Cardon (2012) analytically modeled how FSA contributions could reduce optimal deductible choices. As I show in the next section, using HSA funds as supplemental coverage can increase or decrease deductible choices depending on risk aversion and HSA balances, because HSA funds do not expire, unlike with FSAs. Peter et al. (2015) theoretically analyzed HSA saving and utilization choices but did not focus on HSAs as a tool to reduce moral hazard.

1.3. A Model of Insurance Choice and HSA Saving

1.3.1. The standard model of insurance plan choice

In the U.S., employers can elect to offer health insurance plans to their labor force. Given that menu, employees select a health insurance contract that pays a portion of health care costs in return for an insurance premium. The simplest form of contract provides for full coverage once the employee has paid a deductible, with a higher insurance premium charged for a lower deductible. These are the plans offered in my setting. In the neoclassical model of insurance choice, employees choose the plan that maximizes expected utility given their risk aversion, claim probability, and marginal tax rate. I follow the plan choice literature in assuming preferences satisfy constant absolute risk aversion (CARA), so that for consumption x, $u(x) = -\frac{1}{\gamma}e^{-\gamma x}$, where γ is the coefficient of absolute risk aversion.[20] For family k in year t, a consumption draw under insurance plan j is specified as:

$$x_{jk} = (y_k - \pi_j)(1 - \tau_k) - OOP_{jk} \tag{1.1}$$

where y_k denotes income, π_j denotes the plan premium, τ_k is the family's marginal tax rate, and OOP_{jk} is an out-of-pocket realization under plan j based on family k's *ex ante* cost distribution F_k. The equation reflects the tax preference for insurance that allow premiums to be paid with pre-tax dollars, while out-of-pocket payments are paid with after-tax dollars. Employees have a discrete choice of J insurance plans and choose the plan that maximizes their expected utility:

$$V_{jk} = \int_0^\infty -\frac{1}{\gamma}\exp\left(-\gamma x_{jk}\right) dF_k\left(OOP_{jk}\right) \tag{1.2}$$

For a given cost distribution F and marginal tax rate τ, the choice of plan identifies a set of risk aversion for each employee consistent with expected utility maximization, because there is a finite menu of insurance contracts to choose from. If employees could instead choose from a continuous menu of insurance contracts, then the choice of deductible would point-identify risk aversion for each employee conditional on his cost distribution and tax rate.

[20] See e.g. Einav et al. (2013); Handel (2013); Handel and Kolstad (2015).

1.3.2. Insurance choice with HSAs

HSA contributions can represent saving for short-term health care costs or saving to finance any form of consumption (including health care) in the future. If HSAs are used to pay for current health care costs, the chosen contribution will be related to the choice of insurance deductible and the size of existing HSA balances. If, instead, HSA contributions represent saving for future consumption, there will be no link between HSA saving and deductible choices because no HSA funds will be withdrawn regardless of current out-of-pocket payments incurred. In this case, employees would treat their HSA like a 401(k) for retirement consumption. This section makes these ideas precise and incorporates HSA saving into the standard model of insurance choice.

First consider HSA contributions for short-term health care costs. HSAs then serve as a way to reduce the deductible. In effect, HSAs provide a continuous choice of insurance contract by allowing people to supplement their existing coverage through tax-preferred contributions.[21] Specifically, people can purchase a dollar of additional coverage at a premium equal to 1 minus their marginal tax rate. In this way, HSA contributions to reduce the deductible function as "deductible insurance." Figure 1.1 below displays the schedule of out-of-pocket payments versus medical expenses for a \$3,000 deductible and different levels of HSA contributions. With no HSA contribution, the insurance contract is represented by the solid black line where the employee pays for all costs until \$3,000, and the plan pays for any additional costs. If used as deductible insurance, HSA contributions shift the benefit schedule down towards the horizontal axis. Any level set within the shaded area is now a feasible contract. For example, the dashed line denotes an HSA contribution of \$1,000, so that the first \$1,000 of costs are paid with pre-tax HSA funds, the next \$2,000 of costs are

[21] The intuition is similar to Medigap insurance for Medicare's out-of-pocket payments.

paid from after-tax funds, and costs beyond $3,000 are paid by the insurance plan.

When used to reduce the current deductible, HSA contributions h now represent a second insurance premium by making the consumption draw:

$$x_{jk}(h|\pi,\tau) = (y_k - \pi_j - h_{jk})(1-\tau_k) - OOP_{jk} + \underbrace{\min\{OOP_{jk}, h_{jk}\}}_{\text{HSA withdrawal}} \tag{1.3}$$

Each OOP realization may now be financed with pre-tax funds withdrawn from the HSA, represented by the last term. This formulation nests the standard model represented in equation (1) as a special case when $h = 0$.

In my setting, the employer matches the first L dollars of employee contributions according to a rate m, where $m = 1$ denotes a dollar-for-dollar match. For a given HSA contribution, a higher match rate provides for additional coverage, so that a consumption draw is then represented as:

$$\begin{aligned} x_{jk}(h|\pi,\tau,m,L) &= (y_k - \pi_j - h_{jk})(1-\tau_k) - OOP_k \\ &\quad + \min\{OOP_k, \min[h_{jk}(1+m_k), h_{jk} + m_k L_k]\} \end{aligned} \tag{1.4}$$

The consumer then chooses the insurance deductible and HSA contribution that maximizes expected utility.

When used to reduce the deductible, HSAs are treated like FSAs. There is additionally some option value from HSA contributions that roll over if unused, unlike FSA contributions. Yet this option value is likely small because consumers can make additional HSA contributions over and above their payroll deductions. So a consumer planning to use his HSA to finance short-term costs could make a small payroll contribution and then contribute more upon incurring medical expenses. In addition, there is little option value from building up HSA assets to hedge against future deductible increases because the deductible is never larger than the statutory maximum HSA contribution in this setting. Consequently, an employee is never at risk for paying more out-of-pocket than he could contribute to his HSA that same year. Modeling insurance plan choices in a static framework therefore does not omit an important continuation value term that would be present in a dynamic model.

If contributions are used only to save for future consumption, there is no relationship between HSA saving and current out-of-pocket costs because no funds will be withdrawn today. There would therefore also be no link between deductible choices and HSA saving. HSAs will be more effective at reducing moral hazard if they are used as a long-term savings vehicle, because consumers would trade off health care versus other consumption.

To quantitatively measure this effect, I introduce a parameter η as the proportion of HSA funding allocated to financing current health care costs. This MPC parameter allows consumers to save for both short-term costs and future consumption, with $\eta = 1$ indicating all funding allocated to current costs, $\eta = 0$ indicating all funding allocated to future costs, and an intermediate value indicating a mixture. The higher is η, the more the HSA functions as a price reduction for current health care costs. Denoting HSA balances for family k as H_k, a consumption draw is then represented as

$$\begin{aligned} x_{jk} &= (y_k - \pi_j - h_{jk})(1 - \tau_k) - OOP_k \\ &\quad + \min\left\{OOP_k, \eta\left(H_k + \min\left[h_{jk}(1 + m_k), h_{jk} + m_k L_k\right]\right)\right\} \end{aligned} \tag{1.5}$$

Risk aversion and the MPC are the two parameters we need to estimate in this model. These parameters do not enter additively, because consumption, which is a function of the MPC, is multiplied by risk aversion. Separately identifying the two parameters requires exploiting sources of variation that independently affect risk aversion but not the MPC, and vice versa. For example, variation in premiums, which affects the demand for different deductibles but not for HSA saving, identifies risk aversion separately from the MPC. Variation in 401(k) saving identifies the MPC separately from risk aversion: when used to save for retirement, HSA saving will be related to 401(k) saving, but 401(k) saving should be unrelated to deductible choices. Other sources of price variation, including the deductibles offered, the employer's matching rates for HSA contributions and the employee's marginal tax rate, influence both plan choice and the level of HSA saving and provide additional sources of identification. In my setting, the employer fully replaced its traditional insurance with HDHPs and HSAs and it then varied the price of its deductibles and matching rates over time and across employees. The empirical variation and details of the administrative data are presented in Section 4. Identification and estimation of the model is presented in Section 6. Before explaining the identifying variation and setting background in detail, I first graphically display the implications of HSA saving decisions on insurance plan choice.

1.3.3. An illustration of the HSA model on plan choices

If used to pay for current health care costs, HSAs can induce different insurance plan choices compared to the standard model. The intuition is that HSAs reduce the financial risk from uncertain consumption associated with paying out-of-pocket payments below the deductible. HSAs can induce different choices in two ways. First, people planning to finance out-of-pocket costs from a large HSA balance may choose a higher deductible than without an HSA, since the higher deductible comes with a lower premium. Second, if HSA balances are small or zero, a person may choose a lower deductible than without an HSA. In effect, HSA contributions function as self-insurance against the deductible. Although he would then pay higher premiums, he could experience utility gains from consumption smoothing under sufficiently high risk aversion. These examples are now illustrated graphically.

Figure 2.1 displays the deductible chosen based on the HSA plan choice model, assuming a coefficient of absolute risk aversion γ of 2.5×10^{-4}.[22] The figure plots choices for the highest and lowest cost distributions for females aged 35-44 using the 2008 menu of premiums and deductibles. The size of HSA balances are plotted on the horizontal axis and the MPC is plotted on the vertical axis. The predictions of the standard model correspond to $\eta = 0$. For the high cost distribution (Panel A), the standard model predicts a deductible choice of \$1,250. For $\eta > 0$, though, different combinations of balances and η can induce higher deductible choices. For the median cost distribution (Panel B), the \$1,750 deductible is chosen in the standard model where $\eta = 0$. If $\eta > 0$, however, then the size of balances can influence which plan is chosen. If balances are zero and η exceeds 0.15, then the lowest deductible plan will be chosen instead of the second-lowest deductible. If, instead, η is high but balances are large, then one of the two highest deductible plans will be chosen.

[22]This level of risk aversion is consistent with recent estimates from the literature that separate risk preferences from information frictions or beliefs about expenditure risk Handel and Kolstad (2015). Earlier studies estimate higher levels of risk aversion while acknowledging that econometric models that do not model all elements of the decision process often rationalize choices with implausibly high risk aversion (see e.g. Sydnor 2010).

1.4. Setting and Data

The company is a top-5 health insurer by both market share and revenues in the U.S. with employees throughout the country. In terms of representativeness, the average salary, age, and tenure among the company's employees are roughly in line with U.S. labor force averages, but about 70 percent of employees are female, substantially higher than many other settings. The clean variation in a menu of insurance deductibles and prices makes this setting particularly important for studying insurance and saving choices. Between 2005 and 2011, the company made changes to its health insurance and retirement saving programs for its own 20,000 employees. It fully replaced its standard insurance contracts with HDHPs in 2008, introducing a menu of HDHP offerings that were differentiated only in the size of the deductible and were otherwise equivalent (e.g. provider networks were identical). I study saving and insurance decisions of the company's employees in 2008 and 2009 to isolate variation in policies related to the company's HSA program.[23]

In 2005, the company began offering employees the choice of HDHPs in addition to traditional health insurance plans. Both types of plans featured relatively simple benefit designs and had identical provider networks. The traditional insurance plans offered first-dollar coverage for physician office visits and prescription drugs. Patients paid $25 or $40 copayments for office visits and received an allowance for each drug fill that varied by tier of the drug. Preventive care—immunizations, physical exams, and certain cancer screening—was free of charge. Patients paid 100 percent out-of-pocket for hospital care and other services until they reached their deductibles. Beyond the

[23] Prior to 2007, U.S. law stipulated that HSA contributions could not exceed the chosen insurance plan's deductible, which biases contribution and deductible choices to be positively correlated. This law was removed in December 2006, but I exclude employees who enrolled in the HSA prior to 2008 since they may have been unaware of the rule change.

deductible, the plan paid for all costs, excluding copayments for office visits which the patient still paid each visit. In the HDHP plan, preventive care was also free and the patient paid the full charge for all other care (physician visits, drugs, hospital care, etc.) until the deductible had been met. After the deductible was met, the plan paid for all charges. In 2008, the company dropped its traditional insurance plans and only offered HDHP plans to its employees. Employees had to make an active choice of insurance plan each year during open enrollment. There was no default option.

The menu of HDHP deductibles ranged from \$1,250 to \$3,000 for employee-only coverage and double that for family coverage. Plans differed only in their premium and size of the deductible, both of which varied over time. The employer contributed a flat amount to each plan premium and then required employees to pay higher costs of additional coverage. Figure 3.1 plots the benefits (spending less premiums and out-of-pocket payments) for insurance plans in 2008 for self-coverage.

Only employees who chose an HDHP plan could open an HSA. There is no default employee contribution to the HSA, unlike with the 401(k). Employees are immediately vested for both their contributions and the employer's contributions. The employer matches employee HSA contributions at different rates and up to different limits based on employees' salary levels over time. Matching rates by year and salary levels are shown in Table 1.2. The price of a dollar contributed by the employee is $\frac{1}{1+m}$, where m is the proportion matched by the employer.

The HSA operates as both a low-interest rate savings account and an investment account. Once balances reach \$2,000, HSA assets can be invested in a variety of mutual funds provided the amount invested exceeds \$1,000 and remains above this level. There are no initial setup fees or monthly fees for the HSA account or the investment account paid by the employee.[24] Low fees and the ability to invest in mutual funds are important, because they rule out the argument that consumers should rationally not use their HSA for retirement saving if investment opportunities are poor and transaction costs are high. Employees are given a debit card to make withdrawals from the non-investment portion of their HSA, which reduces the transaction costs of using HSA funds compared to filing paperwork.

The company pursued an extensive communications campaign to inform employees about the HDHP offerings and HSA program. This effort included materials and programs to aid employees in analyzing insurance options and monitoring expenditures. Employees received an annual "Smart Summary" with details on their spending patterns and indicating alternative plans or therapies that may save costs. The employer also provided online budgeting tools, cost calculators, and other resources on their insurance and saving products.

In terms of retirement benefits, the company offered employees a defined-contribution 401(k) and matched employee contributions up to 6 percent of salary. Prior to 2008, the company matched all employee contributions at 50 percent up to this threshold. Starting in 2008, the company began matching the first percent of employee salary 100 percent and then matched subsequent contributions at 50 percent, up to 6 percent of salary (so that the maximum employer contribution increased from 3 to 3.5 percent of salary). Employee contributions were deducted from each period's paycheck. Unlike with the HSA, the employer provided a higher match limit in absolute dollars to higher salaried workers because the 401(k) match was based on a percentage of salary. If employees did not actively enroll in the 401(k) when they were hired, they were auto-enrolled at a salary contribution of 4 percent, also deducted from each period's

[24]The investment account charges a $20 transaction fee when funds are purchased and sold and a $25 fee if there is no activity for 12 continuous months.

paycheck.[25] 401(k) accounts were charged a periodic administrative fee equal to a variable percentage of the balance.

The administrative data includes detailed information on each employee's salary, job characteristics, demographics, medical and pharmacy claims, and choices about retirement saving and health insurance plans. Information on job characteristics and geographic location is measured once in 2011, while information on salary, retirement saving and health insurance choices, and medical and pharmacy claims are measured repeatedly over time. 401(k) contributions (both by the employee and employer) and balances are measured annually. HSA variables—contributions by both the employee and employer, employee withdrawals, balances, and interest—are measured monthly. I aggregate HSA contributions and balances to annual levels to accord with the 401(k) data. Insurance choices are measured annually, and employees could not switch plans during the year.

The claims data includes information on health expenditure for employees and any dependents covered under the employee's policy. Each claim provides detailed information on diagnoses (ICD-9 codes for medical claims and NDC codes for pharmacy claims), providers, and payment (e.g. patient paid, plan paid), and dates of payment. The employer developed its own estimate of the employee's health expenditure risk, which is captured in a variable called the "severity score."

Using the severity score and the medical and pharmacy claims, I construct distributions of expenditure risk for each employee and his or her dependents. This procedure follows the approach of other studies of insurance plan choices (Handel 2013; Handel and Kolstad 2015). As an overview, the strategy is to group people with similar health

[25] In 2011, the company introduced an "auto-escalation" policy that increased the employee's 401(k) contribution from 4 to 5 percent for those who did not actively enroll upon hiring and did not change their contribution in the following two years.

risks in year t together and use the distribution of actual expenditures from year $t+1$ to generate beliefs about expenditure risk at the time the insurance plan was chosen in year t. Each person in the same risk group is assumed to have the same beliefs about his health risk. More specifically, I first group each insured individual into 60 different risk groups based on his age, sex, and severity score. I do not construct separate risk groups by year due to sample size limitations. For each of the 60 risk groups, I record the empirical proportion of individuals with zero expenditures the following year. For those with positive expenditures the following year, I fit a Weibull distribution to the observed expenditure, estimating the shape and scale parameters of this distribution. I exclude expenditures on preventive care since it was covered free of charge by all plans. I also only consider claims from in-network providers, which comprise nearly all spending. For each risk group, I then construct a "modified" Weibull distribution using the group's estimated shape and scale parameters for positive expenditure and the empirical probability of zero expenditures. For each risk group, I take 100 draws from their corresponding modified Weibull distribution. Within each family, I sum the expenditures for each draw so that each family has 100 draws corresponding to the sum of expenditures for each of its members from that particular draw. This statistical object represents the family's beliefs about its total expenditure risk. The family's out-of-pocket risk is then constructed by applying the insurance plan's deductible to each expenditure draw. In this setting, out-of-pocket payments simply equal spending if below the deductible, and the deductible if spending exceeds the deductible.

1.4.1. Sample composition

The analysis sample is constructed by starting with all employees appearing in the employer's Human Resource records in plan years 2008 or 2009 and taking advantage of the key variation in the data related to the HSA program. I restrict the sample to those who are (i) enrolled in one of the company's health insurance plans, (ii) did not switch the number of covered dependents during the year on their insurance plan, (iii) had coverage the entire year when insured, (iv) participated in the HSA program after 2008 and did not have an HSA prior to 2008, (v) actively enrolled in the 401(k), and (vi) were younger than age 59. Restrictions (i) - (iii) are simply to isolate those whose insurance status is not fragmented (16,636 employees). I exclude employees who opened an HSA prior to 2008 because although the law tying the maximum contribution to the lesser of the deductible or the statutory limit was repealed in December 2006, some employees with an HSA may not have been aware of this rule change. The contribution decisions of employees who did not open an HSA prior to 2008 are less likely to be biased by the old rules governing plan contributions. I exclude the 20 percent of remaining employees who passively default into saving 4 percent in their 401(k) because my economic model of HSA saving views 401(k) saving as a reflection of inter-temporal preferences and retirement saving objectives. This approach requires employees make an active decision of how much to save in their 401(k), rather than be auto-enrolled in the 401(k). I finally exclude the small number of remaining employees aged 59 years and older because 401(k) assets can be withdrawn penalty free for any reason starting at age $59\frac{1}{2}$. After these restrictions, the sample size totals 5,314 employees.

Table 1.3 presents summary statistics of the analysis sample, overall and by type of insurance coverage. The average age is 39 years, the average tenure with the firm is 6.4 years, and the average salary is $54,517. Seventy-one percent of the sample is female. Annual HSA saving, including employer contributions, averages $1,279 for self-only coverage and $2,410 for family coverage. Over 96 percent of the sample receives the full employer HSA match and 3 percent contribute the maximum to their HSAs. By way of comparison, 43 percent receive the full employer 401(k) match and just under 2 percent contribute the maximum to their 401(k)s. The lower contribution thresholds for the HSA explain the higher share of employees obtaining the full HSA match. For each account, the large majority of employees are not at a corner solution of contributing zero or the maximum amount. Comparing withdrawals to contributions, nearly 80 percent of funds are withdrawn during this two-year period. On a per capita basis (rather than dollar-weighted), the average share of funds withdrawn is around 70 percent. That statistic points to consumers using HSAs to save for current medical costs, rather than future consumption.

The menu of premiums, deductibles, and enrollment by plan is presented in Table 1.4. For both self and family coverage, the lowest deductible plan was most popular in both years, with the highest deductible as the second-most popular plan. The lowest market shares for the two middle plans is consistent with the benefit schedule plotted in Figure 3.1. Importantly, nearly 30 percent of consumers switch between the four deductible tiers over two years, with roughly equal numbers increasing and lowering their deductibles. This high switching rate is remarkable, considering that no plan was dominated, in contrast to other research documenting lower switching rates in settings with dominated plans (Handel 2013, Bhargava, Loewenstein, Sydnor 2015).[26]

[26]It is worth noting that even before HDHPs were introduced, the rate of switching was 26 percent, still very high compared to other contexts. During this time, employees had access to a non-portable, employer-funded spending account that enabled unused dollars to roll over—similar to an HSA without the retirement savings option. Experience with this account may explain the observed switching rates prior to HDHPs and HSAs.

1.5. Descriptive Evidence

Before estimating the plan choice model, this section presents descriptive regressions of plan and saving choices showing that consumers use HSAs to finance short-term health care costs. These regressions include other endogenous choices (either deductible or saving choices) as explanatory variables and so should not be interpreted as causal, but two key correlations provide evidence supporting the argument that HSAs are not used solely to finance future consumption. First, higher HSA balances at the start of the year are negatively correlated with subsequent HSA contributions. As a placebo test, there is no relationship between 401(k) balances and subsequent 401(k) saving, which is expected since 401(k)s are largely used for retirement saving. Second, higher HSA balances are correlated with higher deductible choices. In addition, there is no evidence that the introduction of HSAs crowds out 401(k) saving, consistent with using the HSA to finance current health care costs. I also test whether the MPC equals 1 and whether the MPC equals 0 using data on HSA saving and withdrawals, rejecting the null in both cases. Finally, I show evidence that health care spending did not decline after traditional coverage was replaced with HDHPs and HSAs.

1.5.1. Reduced-form regressions of HSA and 401(k) saving

Table 1.5 presents the results OLS regressions of employee HSA saving with employee and year fixed effects. The regressions take the following form:

$$y_{kt} = \alpha + \beta_1 H_{kt} + \beta_2 D_{kt} + \beta_3 p_{kt}^{HSA} + \beta_4 K_{kt} + \beta_5 p_{kt}^{401k} + X\delta + \varsigma_k + \upsilon_t + \varepsilon_{kt} \quad (1.6)$$

where y_{kt} denotes employee HSA saving of employee k in year t, H denotes HSA balances and K denotes 401(k) balances at the beginning of year t, p_{HSA} denotes the price of HSA contributions and p_{401k} the price of 401(k) contributions—defined as $\frac{1}{1+m}$ where m is the corresponding matching rate, D_{kt} represents the (endogenous) chosen deductible, ς_k denotes employee fixed effects, υ_t denotes time fixed effects, and ε_{kt} represents a normally distributed random error. The vector of controls X includes expected health spending, a cubic in age, indicators for salary, job tenure, number of dependents, and an indicator for family coverage.[27]

The key coefficients of interest are β_1 and β_2. A negative sign on $\hat{\beta}_1$ would indicate that future HSA contributions are offset against existing balances, and point to HSAs being used for short-term costs. A positive sign on $\hat{\beta}_2$ would also suggest that HSA contributions are designed for current consumption. In one specification, I drop the deductible and HSA balance variables and instead include expected out-of-pocket payments less existing HSA balances, which combines both variables using each employee's estimated distribution of out-of-pocket costs. In placebo regressions, I estimate models of 401(k) contributions as the dependent variable to test whether $\hat{\beta}_4 = 0$, which would be consistent with 401(k)s being treated as retirement saving.

In regressions of HSA saving, the coefficient estimate on beginning-year HSA balances is negative, large in magnitude, and highly significant (Table 1.5, columns 1-3). A dollar increase in HSA balances is associated with a 30 cent reduction in that year's HSA contribution. Measured in logs, a doubling of the HSA balance is associated with a 5.9 percent reduction in that year's HSA saving. Since the company's system prompts employees to choose a dollar amount of HSA saving (rather than a percentage of salary as with the 401(k)) and because the deductible is also measured in dollars,

[27]I include the level of current 401(k) saving rather than its lag in order to capture the effect of contemporaneous shocks that could lead to revising saving plans from one year to the next. I later document that HSA saving does not appear to crowd-out 401(k) saving.

the levels interpretation is a more natural representation of the employee's choice problem. Nonetheless, the key qualitative result—that higher starting balances offset contributions—is similar regardless of functional form, which suggests most people treat HSAs to finance current out-of-pocket costs. The estimated positive sign on expected OOP risk less HSA balances (Column 2) also points to this conclusion. The estimates from placebo regressions of 401(k) saving in Columns 4 and 5 of Table 1.5 demonstrate beginning-year 401(k) balances do not have a strong association with that year's 401(k) contribution. This weak correlation is expected since 401(k)s are designed as long-term saving vehicles.

1.5.2. Reduced-form regressions of plan choices

The second set of descriptive regressions study how HSA balances and saving relate to plan choice. I run fixed effects logit regressions that specify the probability of choosing the highest deductible as:

$$\Pr(D_{kt}^{HIGH} = 1 | \varsigma_k, \beta) = \frac{\exp(\varsigma_k + x_{kt}\beta)}{1 + \exp(\varsigma_k + x_{kt}\beta)} \tag{1.7}$$

where D_{kt}^{HIGH} is a binary variable equal to 1 if the highest deductible is chosen and 0 otherwise, ς_k is a an employee fixed effect, and and x_{kt} represents the vector of regressors that include HSA and 401(k) balances and contributions, expected health risk, and other controls appearing in Table 1.5. I also run the same regression with the lowest deductible chosen as the dependent variable. The first two columns of Table 6 present odds ratios of choosing the lowest and highest deductibles, respectively. Since the estimation only uses information from observations where the dependent variable changes over time, I combine self and family coverage to increase sample size. Based on the estimates, a one-standard deviation increase in the HSA balance at the beginning of the year is associated with a 15 percent reduction in the odds of choosing the lowest deductible and a 10 percent increase in the odds of choosing the highest deductible. By way of comparison, a one standard deviation increase in expected health spending increases the odds of choosing the lowest deductible by 42 percent and reduces the odds of choosing the highest deductible by 40 percent. Using deciles of expected health spending yields similar qualitative results. Other saving variables are not statistically significant predictors of plan choice.

Since plans differ only in the size of their deductibles, I also run ordered logit models with the deductible as the dependent variable. The regressions now exploit between-employee variation and include indicators for being married and white in addition to the controls included in Columns 1 and 2. The coefficient estimates for self and family coverage are presented separately in Columns 3 and 4 of Table 1.6. Higher HSA balances are again correlated with higher deductibles.

To gauge the magnitude of the ordered logit estimates on HSA balances, I calculate the probability that each plan in 2009 is selected at different levels of beginning-year HSA balances, holding all other variables at their means. As shown in Table 1.7, moving from an HSA balance of zero to the sample's mean balance is associated with a roughly 3 percentage point decrease in the probability of choosing the lowest deductible and a 2 percentage point increase in choosing the highest deductible. Moving from a zero balance to the 95th percentile lowers the probability of choosing the lowest deductible by almost 10 percentage points and increases the probability of choosing the highest deductible by roughly 8 percentage points for both coverage types The magnitude of these changes in deductible choices is large, similar to results from the fixed effects logit.

1.5.3. *Crowd-out of 401(k) saving*

There is little evidence that 401(k) saving declines after HSAs are introduced. The apparent absence of 401(k) crowd-out is consistent with using HSAs to finance short-term health care costs. In results not shown, there is no statistically significant relationship between total tax-preferred saving—defined as the sum of 401(k) and HSA saving—and the introduction of HSAs. Nevertheless, HSAs raise the interest rate on saving, so this weak correlation could result from offsetting income and substitution effects. The unavailability of data on other assets prevents me from examining this question.[28] Yet the observation that 401(k) saving did not drop after the introduction of HSAs is consistent with the use of HSAs as a short-term savings vehicle.

1.5.4. *Tests of MPC=0 and MPC=1*

Before estimating the MPC using a structural model of insurance plan choice, it is instructive to test the two extreme cases that HSA funds are treated like 401(k)s—in which the MPC equals 0—or are treated like FSAs—in which the MPC equals 1. One approach to test these hypotheses is to compare HSA withdrawals to out-of-pocket costs and HSA balances. Define M as the minimum of out-of-pocket costs and HSA assets (beginning-year balances plus current contributions). When HSA assets exceed out-of-pocket costs, HSA withdrawals should only be compared to the incurred out-of-pocket payments. When HSA assets are insufficient to cover expenses, then HSA withdrawals should only be compared to available HSA funds. Withdrawals should

[28] The relationship between HSA and 401(k) funding relates to the long debate on retirement saving crowd-out and the interest elasticity of saving. Early studies on the introduction of IRAs and 401(k)s such as Engen et al. (1996), Poterba et al. (1996), and Benjamin (2003) reached different conclusions partly because they struggled to isolate exogenous variation in 401(k) contributions from tastes for saving, as reviewed by Bernheim (2002). The most comprehensive study to date (Chetty et al. 2014) using Danish tax registry data estimates that 99 cents of every dollar saved in a tax-preferred account would have been saved in non-taxable accounts, indicating almost full crowd-out.

equal M if the MPC equals 1 and should equal 0 if the MPC is 0. I regress the level of HSA withdrawals on M and observables, which include expected health spending, salary, the number of dependents, type of coverage, age, sex, and firm tenure. The point estimate on the coefficient on M is 0.733. I test whether the coefficient estimate is either 1 or 0, rejecting the null in both cases. If I include employee fixed effects, I obtain smaller point estimates but similar results of the hypothesis tests.

1.5.5. Changes in Health Care Spending

This section quantifies the change in health care spending associated with the replacement of traditional insurance coverage with HDHPs and HSAs. Studies of HDHPs in other settings have tended to find moderate reductions in health care spending in the range of 10 to 15 percent (Buntin et al. 2011; Borah et al. 2011; Haviland et al. 2015; Brot-Goldberg et al. 2015). By contrast, HDHPs were not linked to a reduction in health care costs in this sample. As described in Section 4, the traditional insurance plans offered before 2008 featured first-dollar coverage, with copayments for physician services and prescription drugs. Employees paid a deductible for other services, which were then fully covered after the deductible. Plans differed only in the size of this deductible. The switch to HDHPs dropped the first-dollar coverage on physicians services and prescription drugs, applying a single deductible for all care (which was also equal to the out-of-pocket (OOP) max, as before).

Table 1.8 presents regressions of the log of total health care spending against an indicator for being enrolled in a HDHP with HSA (corresponding to years 2008 and 2009), salary, expected health spending, the number of dependents, and employee fixed effects. Among all employees in the sample, the introduction of HDHPs was associated with a 6 percent increase as shown in Column 1, which is statistically significant at the 10 percent level. Columns 2 and 3 compare the sub-sample of employees who selected similar coverage levels before and after the introduction of HDHPs, defined as having a standard deviation in the deductible of less than $250 throughout the period. The coefficient estimate on the HDHP and HSA indicator is close to zero and imprecisely estimated. The fourth column restricts the sample to those employees with the same out-of-pocket maximum before and after HDHPs. Total spending is estimated to be lower by 5.3 percent after the introduction of HDHPs with HSAs, but again this estimate is not statistically significant. Replacing traditional coverage with HDHPs and HSAs was not linked to sizable spending declines in this setting.

1.6. Estimation and Identification of Choice Model

1.6.1. Overview

The choice model is estimated by simulated maximum likelihood. The estimating algorithm draws parameter values from the assumed distributions of the random coefficient (risk aversion), calculates choice probabilities for each sequence of choices given those parameters, and then matches predicted and observed choices. The standard model includes risk aversion as the only source of preference heterogeneity. The HSA model also includes the MPC and allows HSA and 401(k) saving to shift demand for insurance plans. Both models incorporate the same cost distributions as described below and construct marginal tax rates using NBER's TAXSIM calculator.[29]

[29] The employee's salary and state of residence are the two inputs to estimate the marginal tax rate using NBER's TAXSIM. Information on salary of spouses and other relevant tax information is not available. I construct the marginal tax rate for employees with family coverage using the employee's salary assuming the taxpayer files as single. This may misestimate the marginal tax rates for a subset of observations, but the bias is likely not systematic as some employees may be in higher brackets and some in lower brackets than assigned based on their salary alone.

1.6.2. Construction of Out-of-Pocket Cost Distributions

This section describes in detail the procedure for constructing distributions of out-of-pocket costs for each insured family (employee only or the employee and dependents). It follows similar methods as Handel (2013) and Handel and Kolstad (2015). This cost model assumes that there is no moral hazard and that each person in the same risk group holds the same beliefs about his or her ex ante health expenditure risk. There are four steps to construct the distributions from the inputs of expenditure claims and the employer's severity score.

1. Group each insured individual i into risk group z based on age, sex, and health status

2. For each risk group, construct a Weibull distribution, G_z, that is modified to allow for the possibility of zero expenditure using observed total health expenditure m from the following year

3. For each person in risk group z, simulate expenditure draws from G_z and add up the draws within each family k to create an ex ante distribution of total health expenditure risk G_k for family k

4. For each family k, map the distribution of expenditure risk G_k to out-of-pocket costs under deductible j to create a family-specific ex ante distribution of out-of-pocket costs F_{jk} of choosing deductible j

Each individual i is first categorized into risk group z based on their age, sex, and quintile of the severity score. The age bins used are 0-14, 15-24, 25-34, 45-64, 65 and older. The severity score, which is recorded on each insurance claim, measures the expected health spending for that enrollee using a proprietary formula constructed by the company. The score is not used for payment purposes, but rather represents the employer's actuarial forecast about that person's expenditure risk. I use the severity score captured on the last claim before the start of the plan year as a measure of health status during the open enrollment period. For each age bin and each sex, I classify individuals into quintiles of the severity score within that age-sex cell. I pool years 2007 through 2009 together to ensure adequate sample sizes. This process results in 60 risk groups based on six age bins, sex, and quintiles of severity score.

After the risk groups are defined, the observed expenditures for each person in the group the following year are used to estimate an ex ante expenditure distribution for that group. Denote the empirical distribution of claims the following year by $\widehat{G}_{I_z}$. In constructing this distribution, expenditures on preventive care are excluded since such services are covered free of charge by all plans. Only claims from in-network providers are considered, which comprise over 95 percent of all spending. I continuously fit this empirical distribution using a Weibull distribution with a mass of claims at zero to generate an ex ante distribution of expenditure risk, consistent with prior work (Handel 2013; Kolstad and Handel 2015).

The creation of this ex ante distribution of expenditure by risk group involves two steps to deal with the mass of expenditure at zero. First, for each risk group k, the empirical probability of zero expenditure is used to construct the mass of expenditure realizations at zero, denoted $\widehat{G}_{I_z}(0)$. Second, a Weibull distribution is fitted to the observed expenditures that are positive in that risk group by maximizing the following likelihood with respect to the scale parameter α and shape parameter β:

$$\prod_{i \in I_z} \frac{\beta_z}{\alpha_z} \left(\frac{m_i}{\alpha_z}\right)^{\beta_z - 1} e^{-\left(\frac{m_i}{\alpha_z}\right)^{\beta_z}}$$

Denote $\widehat{\alpha_z}$ and $\widehat{\beta_z}$ as the estimated parameters and $W\left(\widehat{\alpha_z}, \widehat{\beta_z}\right)$ as the distribution of positive expenditure in risk group z. The (ex ante) distribution for expenditure in risk group z is then:

$$G_z = \begin{cases} \widehat{G}_{I_z}(0) & \text{if } m = 0 \\ \widehat{G}_{I_z}(0) + \frac{W\left(\widehat{\alpha_z}, \widehat{\beta_z}\right)}{1-\widehat{G}_{I_z}(0)} & \text{if } m > 0 \end{cases}$$

For each insured individual within each risk group, 100 draws are simulated from the corresponding expenditure distribution G_z. Then within each family k, the expenditures for each draw from each member are summed, so that each family has 100 draws corresponding to the family's total expenditure. This statistical object, denoted G_k, represents the beliefs of family k about its total health expenditure risk. Since families differ in their compositions by age, sex, severity score, and size, this classification by risk group results in over 2,500 different combinations of expected spending in the sample.

Each family's out-of-pocket cost distribution in plan j, denoted F_{jk}, is calculated by applying each expenditure draw from distribution G_k to the insurance deductible D_j. Since there is no coinsurance beyond the deductible, out-of-pocket costs equal total health expenditure if expenditure is below the deductible and equal the deductible if expenditure exceeds the deductible. There is a single deductible for both medical and pharmacy claims so the out-of-pocket mapping between total expenditure and costs is particularly simple in this setting. The mapping is thus defined as:

$$OOP = \begin{cases} m & \text{if } m \leq D_j \\ D_j & \text{if } m > D_j \end{cases}$$

The family-specific distribution of out-of-pocket costs F_{jk} under each deductible is then used as input into the model of insurance and saving choices.

1.6.3. Specification

Observing 401(k) saving allows me to estimate the model's parameters without estimating a fully structural life-cycle model of savings. I assume any unobservables—such as beliefs about income processes or bequest motives—and preferences for future consumption operate equivalently between the HSAs and 401(k)s given the nearly identical structure and rules between the two savings vehicles. Importantly, this approach does not require assumptions about the optimality of 401(k) saving, which may be violated in practice (Choi et al. 2011).

Risk aversion, which is assumed to be time-invariant, is modeled as a random coefficient and assumed to follow a normal distribution with mean and variance:

$$\gamma \sim N\left(\mu_\gamma, \sigma_\gamma^2\right) \tag{1.8}$$

The mean is specified to be linearly related to employee age, income, job characteristics, and coverage type as:

$$\begin{aligned} \mu_\gamma &= \alpha_\gamma + \beta_\gamma^{age} AGE + + \beta_\gamma^{sal} SALARY + \beta_\gamma^{fin} FINANCE \\ &\quad + \beta_\gamma^{tenure} TENURE + \beta_\gamma^{self} SELF \end{aligned} \tag{1.9}$$

where employee age is specified in 2008, salary is averaged between 2008 and 2009,[30] *FINANCE* is an indicator for whether the employee's job classification relates has a label of Finance/Accounting, Actuarial, or Underwriting, *TENURE* is an indicator for working at the firm at least 6 years, and *SELF* is an indicator for self-coverage.

[30] The specification of CARA utility assumes no income effects, so including salary is intended to pick up the influence of variables correlated with salary, rather than the role of income per se.

Consumption for family k in plan j in year t in the standard model is:

$$x_{jkt} = (y_k - \pi_{jt})(1 - \tau_{kt}) - OOP_{jkt} + \epsilon_{jkt} \tag{1.10}$$

where ϵ_{jkt} denotes a normally distributed family-plan-time-specific error that is iid for each plan. The standard deviation of the errors, denoted $\sigma_{\epsilon j}(C_k)$, is allowed to vary between self and family coverage, indexed by C_k for family k. This error rationalizes choices not predicted given the candidate preference parameters, ex ante health care costs, and other observables. Conceptually, the error may represent errors in forecasting health care costs. The variance of the shock on the lowest deductible plan in each year is normalized to 1.

In the HSA model, the MPC is modeled as a linear function of the same characteristics as:

$$\eta = \alpha_\eta + \beta_\eta^{age} AGE + \beta_\eta^{sal} SALARY + \beta_\eta^{fin} FINANCE + \beta_\eta^{tenure} TENURE + \beta_\eta^{self} SELF \tag{1.11}$$

The coefficient estimates in the above equation provide information about the potential mechanisms behind using the HSA for short-term costs. For example, a negative sign on $\hat{\beta}_\eta^{age}$ would support discounting, because older employees—who have to wait fewer years until HSA funds become unrestricted—would then use HSAs more for future consumption compared to younger employees. Income and job type provide indirect evidence regarding financial literacy, with negative coefficient estimates on those variables supporting that mechanism.

The estimation of the HSA model adds three terms to the standard model as shown in equation below. Consumption in the standard model is represented in the first line. The term in the second line captures the HSA withdrawal to finance a given OOP payment as a function of the MPC η, employee contributions h, employer matching rates m and limits L, and beginning-year HSA balances H. This term captures the interaction between current out-of-pocket risk and HSA funding. The third line incorporates HSA and 401(k) saving as demand shifters for plan choices. Conditional on 401(k) saving, if HSA saving does not differ across plans, then $\widehat{\eta} = 0$. If $\eta > 0$, then HSA saving is chosen partly to offset current out-of-pocket expenditure. In the case where $\eta = 1$, HSA saving fully represents a price reduction for current out-of-pocket payments.

$$\begin{aligned}
\tilde{x}_{jkt} = {} & (y_k - \pi_{jt} - h_{kt})(1 - \tau_{kt}) - OOP_{jkt} \\
& + \min\left\{OOP_{jkt}, \eta\left(H_{kt} + \min\left[h_{kt}(1 + m_{kt}), h_{kt} + m_{kt}L_{kt}\right]\right)\right\} \\
& + (1 - \eta)\sum_{j \in J} h_{kt}\mathbf{I}(d = j) + \sum_{j \in J} \kappa_{jt} 401(\text{k})_{kt}\mathbf{I}(d = j) + \widetilde{\epsilon}_{jkt} \quad (1.12)
\end{aligned}$$

Including HSA and 401(k) saving as demand-shifters is a reduced-form approach to estimate the MPC and risk aversion. The specification, motivated by theory, constrains η to lie between 0 and 1 and estimates η by running a horse race between 401(k) saving—capturing saving for future consumption—and HSA balances, matching rates, and interactions with out-of-pocket payments—capturing saving for current health care costs. A more structural approach would model the continuous choice of HSA saving simultaneously with the discrete choice of insurance deductible in a framework similar to Dubin and McFadden (1984). That formulation would increase the choice set each year from J to JQ, where Q is the number of possible HSA contributions. This increase in dimensionality raises the computational requirements for the estimation substantially even with discretizing the set of HSA contributions into larger bins. In Section 7, I perform a calibration as a robustness check that simultaneously models the joint decision of deductible and saving choices. Without an error term, the parameter estimates can be recovered person by person given the richness of the data variation. That exercise yields results qualitatively similar to the choice model here, providing support to the reduced-form approach to model HSA and insurance choices.

The probability of household k choosing plan j in year t for simulation s, with expected utility denoted V_{jkts}, is calculated using a smoothed accept-reject function with a modified logit transformation (Train 2009).[31] Compared to an accept-reject simulator, this transformation accommodates the case where a small change in parameters does not translate into a change in the discrete choice of plan as well as the case where some observed choices are never optimal. Let θ denote the parameters of the model to be estimated: $\theta \equiv \left(\mu_\gamma, \sigma_\gamma^2, \alpha, \beta, \mu_\eta, \sigma_{\epsilon j}\left(C_k\right)\right)$. Given candidate parameters θ, the smoothed accepted-reject simulator specifies the choice probability of family k choosing plan j in year t in simulation s as:

$$\Pr(j = j^*|\theta)_{kts} = \frac{\left(V_j^*\right)^\kappa}{\sum_{\tilde{j}}\left(\widetilde{V_j}\right)^\kappa} \tag{1.13}$$

where $V_j^* \equiv \frac{\frac{1}{-V_{kj^*ts}}}{\sum_J \frac{1}{-V_{kjts}}}$ and $\widetilde{V_j} \equiv \frac{\frac{1}{-V_{k\tilde{j}ts}}}{\sum_J \frac{1}{-V_{kjts}}}$. The parameter κ is a scale factor used to approximate the indicator function of the accept-reject simulator without introducing numerical difficulties.[32]

[31] The logit transformation of the accept-reject simulator was suggested by McFadden (1989). I follow Handel (2013) in taking the reciprocal of the expected utilities, normalized by the utility of one plan, because CARA utility is negative and already exponential.

[32] I use κ=6 in calculating choices and verify the estimates are not sensitive to the choice of smoothing parameter.

With panel data, the choice probability of sequence of choices for family k across simulations is:

$$P_k(\theta) = \sum_{s \in S} \prod_{t \in T} \prod_{j \in J} \Pr(j = j^* | \theta)_{kts} \tag{1.14}$$

With four plan choices and two years, there are 16 possible choice sequences in the standard model where the only discrete choices is the deductible. Denote the set of all sequences of choices as $W = J^T$, where T=2 with two years of data. The simulated log likelihood function is constructed by summing choice probabilities of observed choices across families:

$$SLL(\theta) = \sum_{k \in K} \sum_{w \in W} d_{kw} \ln(P_k(\theta)) \tag{1.15}$$

where d_{kw} is an indicator function for whether family k chose sequence w. The estimating algorithm uses interior point methods to select the value of θ that maximizes $SLL(\theta)$.

1.6.4. Identification

In the standard model, the choice of deductible identifies sets of risk aversion that maximize expected utility based on the ex ante cost distribution, premiums, and marginal tax rate. Risk aversion is point identified by making a parametric assumption on its distribution. Identification also assumes that consumers have rational beliefs about their ex ante cost distribution, which is a standard assumption made in the literature. In the model with HSA saving, the two preference parameters to estimate are risk aversion and the MPC. The employer's introduction of HDHPs, cross-sectional and time-series variation in deductibles offered, premiums, matching rates, and data on 401(k) choices are key to separately identify the parameters. The employer's switch to HDHPs can be viewed as exogenous because the insurer viewed increased consumer decision-making among its own workforce as a strategy that could help reduce costs for its clients too. The adoption was therefore not based on a perceived demand from employees for this type of insurance product. From the perspective of a single employee, cross-sectional and time-series variation in deductibles, premiums, and matching rates are also clearly exogenous.

Risk aversion is separately identified from the MPC because some variation shifts deductible choices but not HSA saving, while other variation shifts HSA saving but not deductible choices. Premium variation influences the choice of deductible but not the level of HSA saving, which provides identifying variation in risk aversion separately from the MPC. Similarly, 401(k) saving affects the level of HSA contributions for retirement consumption, but does not influence deductible choice. Saving in 401(k) plans thus helps to identify the MPC separately from risk aversion. Employer matching rates influence both deductible and saving choices and provides another source of variation to identify the model's parameters. Matching rates vary mostly cross-sectionally, but they also change over time for 9 percent of the sample. Within-employee changes in HSA balances reflect choices on past contributions and withdrawals that also provide important information on how HSAs are used. The necessary assumption for within-employee variation in HSA balances to credibly identify η and γ is that HSA balances affect saving and plan choices only through the expected out-of-pocket risk. Since there is also little option value from HSA saving as discussed in Section 3, HSA balances should affect current plan choices but not future plan choices.

1.7. Results

This section presents results from the choice model that estimates risk aversion and the MPC. The results demonstrate the importance of incorporating HSA funding decisions into models of insurance choices with HDHPs. The standard model struggles to rationalizes these choices with risk aversion as the sole source of preference heterogeneity. Table 9, Column 1 presents the estimated parameters from the standard model. The model estimates a high level of risk aversion—the mean CARA coefficient is 1.5×10^{-3}—and also large heterogeneity in risk preferences. The estimated variances on the random shocks are very large for the highest deductible plan for both self and family coverage. Such large variances relative to the lowest deductible plan rationalize the high rates of switching between deductibles from one year to the next. Observed plan choices are thus rationalized with large enough shocks specific to each year and plan and highly heterogeneous risk preferences.

Estimates that use HSA and 401(k) contributions as demand shifters is presented in Column 2 of Table 1.9. Accounting for HSA contributions leads to a lower estimate of risk aversion, 3.5×10^{-4}, which is more plausible than the risk aversion estimated in the standard model. To interpret the magnitude of the risk aversion estimates, I follow Cohen and Einav (2007) and subsequent work in calculating the amount Y such that a consumer with the given level of risk aversion would be indifferent between accepting and rejecting a 50-50 gamble to win \$1,000 or lose \$$Y$. In the standard model, Y equals \$383 versus \$739 in the HSA model. To express the results in terms of relative risk aversion (RRA), multiplying by the sample's median after-tax salary yields CRRA estimates of 42 and 10, respectively. The point estimates of risk aversion are statistically different between the two models. Including HSAs substantially improves the model's fit to the data, as indicated by the Likelihood

Ratio test statistic which is statistically significant at the 1 percent level. In both models, risk aversion is still estimated to be highly heterogeneous across consumers as reflected by the large standard deviation relative to the mean. Larger dispersion in estimated risk aversion in this setting compared to others may be driven by greater rates of plan switching.

The MPC is estimated to 0.68, on average, with a 95 percent confidence interval between 0.65 and 0.72. Higher income employees are more likely to use their HSA to save for future consumption rather than current medical costs, as indicated by the negative estimate on the MPC slope coefficient. Older employees are more likely to use the account for short-term costs, as are employees with self-coverage. Conditional on these variables, the relationship between the MPC and job characteristics (tenure or working in a finance-related position) is not statistically significant. I return to discuss the potential underlying sources of the MPC in Section 1.8.

1.7.1. Robustness: Calibration using HSA withdrawals

Since the HDHPs are differentiated only by their deductibles and are equivalent across all other dimensions (e.g. provider networks), it is useful to consider a robustness test that excludes an error term and does not specify the distribution of risk aversion. Without a plan- and time-specific shock, the expected utility model is calibrated by taking advantage of the fact that when used to finance the deductible, HSA saving offers a continuous choice of insurance contract that flexibly recovers the distribution of risk aversion. The level of HSA saving intended for current out-of-pocket payments and choice of deductible then point-identifies risk aversion for each observation, conditional on its cost distribution and marginal tax rate. I numerically solve for the level of HSA saving that maximizes expected utility for each observed combination of cost distribution and marginal tax rate for a given risk aversion level. By performing this calculation over a grid of risk aversion from 10^{-6} to 5×10^{-3} in increments of 10^{-6}, I construct a monotonic mapping from risk aversion to deductible choices, HSA saving for current out-of-pocket costs, cost distributions, and marginal tax rates.

Once the amount of HSA saving reserved for current out-of-pocket payments is quantified, applying this mapping recovers each observation's risk aversion coefficient. I take the level of observed withdrawals each year as the simplest estimate of the quantity of HSA saving reserved for short-term costs. The average share of assets withdrawn each year is 70 percent. On the one hand, this estimate of the MPC will underestimate the level of deductible insurance demanded if a person incurs minimal medical claims. On the other hand, the MPC would be overestimated if HSA withdrawals purchase predictable over-the-counter spending, which would not be included in the cost distributions that are based on insurance claims.[33]

Table 1.10 presents sample statistics of estimated risk aversion from this calibration. Median risk aversion is estimated to be 2.4×10^{-4}, slightly lower than the choice model estimate in Table 9. The correlation coefficient between risk aversion and MPC is 0.21, The results of this calibration are consistent with the estimates from the choice model, providing additional support to the finding that HSAs are largely used to finance short-term health care costs.

1.7.2. Robustness: Estimation of choice model using 1st-year data only

I have argued it is reasonable to use within-employee variation in HSA balances to identify the parameters and to not model plan choices as a dynamic optimization problem. As a check on this assumption, I re-estimate the choice model using data

[33] I obtain quantitatively similar results if I calculate the average withdrawal as a percentage of assets during the two-year period because most people withdraw similar proportions of HSA funds each year.

only from 2008 when beginning-year HSA balances were zero for everyone. Premium variation still identifies risk aversion separately from the MPC and 401(k) saving identifies the MPC separately from risk aversion. The MPC is estimated to be 0.76, which is only slightly larger than the estimates in Table 1.9. This check reinforces the results, and suggests the parameter estimates are not driven by dynamics omitted from the model.

1.8. Mechanisms

The MPC summarizes the potential effect of several economic fundamentals on insurance plan choices, including discounting, mental accounting, information frictions, and financial literacy. This section discusses several of these potential underlying sources. I first present empirical evidence related to information frictions and liquidity constraints, which are two mechanisms that the data allow me to partially investigate. I then discuss how discounting and mental accounting may affect choices, and how such fundamentals relate to financial literacy. Understanding the micro-foundations of these decisions may be important for policy. For example, interventions to encourage desirable choices—whether through information, financial incentives, defaults, etc.—are likely to be more effective if they are targeted at the relevant part of the decision process. I find that employees are well-informed about many parts of insurance coverage, including coverage of preventive care, the tax benefits of HSA saving, and the ability of HSA assets to roll over each year. The data do not allow me to study whether employees understand the fungibility of HSA assets as cash in retirement, however, which is arguably the least well known feature of HSAs. I do find suggestive evidence that liquidity constraints partly explain choices. In particular, the inability to finance the entire deductible may lead employees to choose lower deductibles than they would otherwise prefer.

1.8.1. *Information frictions*

In addition to the fact that the company is a health insurer, several pieces of empirical evidence point to the company's employees being more knowledgable about health insurance products than the average consumer. First, employees choose dominated plans less often than has been documented elsewhere. In 2007, before the mandatory switch to HDHP coverage, the lowest deductible plan was dominated by higher deductible plans for family coverage. More specifically, to reduce the deductible from \$4,000 to \$2,500 cost over \$1,500 in higher premiums. For any level or type of spending, a family would therefore pay more if enrolled in the lowest deductible plan than plans with higher deductibles. The lowest deductible plan was not dominated for plans covering an employee only, employee plus spouse, or employee plus one child. The plan shares by coverage type reflects these price differences: the lowest deductible was chosen by 17 percent of consumers with employee-only coverage, employee plus child coverage, and employee plus spouse coverage (where it was not dominated), while it was chosen by only 11 percent with family coverage.[34] By contrast, a recent study where employees of a large firm could customize their plan design found that 55 percent selected dominated plans (Bhargava, Loewenstein, and Sydnor 2015), considerably more than this setting.

Second, HSA saving is lower in states that do not exempt contributions from state income taxes, consistent with employees being knowledgeable about HSA tax benefits. A handful of States—Alabama, California, Pennsylvania, New Jersey, and Wisconsin—did not exempt HSA funds from state income taxes during this period.

[34] It is worth noting that depending on a family's marginal tax rate, this plan may not be dominated because premium payments are tax exempt while out-of-pocket payments are not. This comparison can then be interpreted as reflecting worse terms of trade for the lowest deductible plan relative to other deductibles for families, but not for other coverage types. Regardless, this comparison of plan shares provides indirect evidence that employees respond to prices in insurance plan choices.

According to the employer, there were even complaints by employees working in these states because they did not receive the same tax benefits. Conditional on observables, the level of annual HSA saving is 7 percent lower in these states compared to those that exempt HSA funds from state income tax. This difference is statistically significant and roughly in line with state-level income tax rates, providing further evidence that employees are aware of the tax benefits from HSA funds.

Third, employees appear to understand that HSA funds do not expire at year's end because withdrawals do not spike during the final months of coverage. If employees believed that HSA funds did not roll over (as with FSAs), then HSA assets should be exhausted at the end of each year and one would expect to see higher withdrawals at the end of the year. This pattern is not observed in the data. Withdrawals are actually smallest at the end of year, and this is not due to HSA balances being lower: the share of HSA assets withdrawn is also lower in later months than earlier months. These patterns provide suggest that employees understand that unused HSA funds roll over.

Fourth, preventive care—which is exempt from the deductible—does not decline after the adoption of HDHPs, suggesting employees are knowledgable about HDHP benefits. There is no decline in the number of preventive visits, total spending on preventive care, or the number of screening mammographies after enrolling in an HDHP. In fact, the averages all increase slightly. This behavior contrasts with research from other settings documenting that consumers were unaware that preventive care was covered in HDHPs and many cut back on preventive care once enrolled in a HDHP (Reed et al. 2012; Brot-Goldberg et al. 2015).

Finally, HSA saving does not vary systematically by firm tenure or date of firm exit, which suggests unobserved information may not be important. For example, employees planning to leave the firm soon after the HSA program began may have rationally chosen to contribute less to their HSA if they were uncertain of HSA availability at a future employer. It is reasonable to believe that longer firm tenure is associated with being more informed, conditional on salary, age, and other factors. Compared to employees with two or more years of firm tenure, employees with 1 year of tenure have lower than average levels of HSA saving for self coverage, but not for family coverage. Those who have worked at the firm the longest tend to make higher than average HSA contributions, conditional on income and other observables, but the differences are small in magnitude.

1.8.2. Liquidity constraints

This section evaluates whether liquidity (or cash) constraints, which are unobserved and outside of my model, may explain plan and saving choices in this setting. Choosing an insurance plan entails a financial trade-off between the upfront cost of the insurance premium and the out-of-pocket costs from utilization. Cash constraints may prevent people from purchasing the plan they want because of either high premiums (for plans with lower deductibles) or high out-of-pocket costs (for plans with higher deductibles). I study cash constraints in two ways. First, I model plan choices when employees face a fixed budget for premiums and out-of-pocket payments, and analyze how plan choices change as this budget is relaxed.[35] Second, I compare observed plan choices and saving decisions for employees who take loans from their 401(k)—roughly one-third of the sample—to those who do not. Research on retirement saving finds that those with 401(k) loans are more likely to be liquidity constrained compared to

[35]In this setting, either the lowest deductible or highest deductible always is the lowest cost, on average, at any given level of spending as shown previously in Figure 3.1. Of course, once risk aversion is introduced, the least expensive plan is no longer necessarily optimal because of curvature in the utility function.

those without a loan (Lu and Mitchell 2010).[36] The analysis suggests liquidity constraints may influence plan choices and HSA saving in this setting, but this mechanism only explains a small portion of the empirical patterns.

First consider plan choices when consumers have a fixed amount of money to spend on premiums and out-of-pocket payments. Depending on the size of the cash constraint, the most preferred plan may not be affordable and so a different plan is chosen instead ("the constrained plan choice"). The data does not allow me to observe the size of any employee's cash constraint. To investigate the potential importance of such constraints in plan choice, this exercise compares how different constrained plan choices are likely to be from unconstrained plan choices. If the two choices are not different, then omitting liquidity constraints from the plan choice model is likely not problematic. I consider two statistics of out-of-pocket costs when defining cash constraints. First, a plan is unaffordable if the sum of premiums and expected out-of-pockets exceeds the cash constraint. The second definition bases affordability on the sum of premiums and the maximum out-of-pocket realization from the cost distribution (which is always the deductible because the distribution is continuous). One could make arguments for using either definition, but requiring the employee to be able to finance whatever spending draw occurs, rather than just the expectation of costs, seems to come closer to news reports about HDHP affordability and consumer decisions (Pear 2015).

Figure 4.1 presents an area plot of the regions for which each deductible choice is optimal, given the constraint that out-of-pocket payments and premiums must not exceed the specified cash constraint for health care. As one moves right along the horizontal axis, the cash constraint is relaxed, and as one moves up on the vertical

[36] The authors use income and non-retirement financial wealth to judge liquidity constraints among 2.3 million Vanguard account holders across nearly 1,000 employers.

axis, risk aversion is increased. I consider women aged 35 to 44 in the median quintile of severity score, representing an intermediate level of spending.[37] When cash constraints are based on premiums and expected out-of-pocket costs (the top panel), the expected cost differences between plans are small, represented by the distance between points A and B. At a risk aversion level of 6×10^{-4}, the employee would prefer to choose the \$1,250 deductible if she were not cash constrained. However, that plan has the highest premiums and expected out-of-pocket payments, and so if her budget were below \$2,390, the plan would be unaffordable. As the cash constraint is tightened (moving to the left horizontally), she would choose a plan with a higher deductible. If her budget for health care were below \$2,288—the premium and expected out-of-pocket payment of the lowest deductible plan—then she will not purchase coverage at all (and will not appear in my sample). Since the range between these two points is only about \$100, liquidity constraints are unlikely to explain plan choices when consumers budget for average out-of-pocket payments.

Liquidity constraints are more likely to distort plan choices if cash constraints are based on the sum of premiums and the deductible (bottom panel). Now at low levels of risk aversion, employees who are unable to finance the entire deductible in case of large health expenses would choose a plan that offers a lower deductible in return for a higher premium. The range over which cash constraints would lead people to make a different plan choice is now larger, equal to \$946 and again represented between points A and B. In this case, ignoring liquidity constraints may overestimate risk aversion: for a given cost distribution, the model judges a person choosing a low deductible to be more risk averse than if that person chose a high deductible, when this choice may instead reflect liquidity constraints.[38] The reverse was true in the

[37] I find similar patterns for other cost distributions as well.

[38] The effect of not modeling liquidity constraints on risk aversion here is similar to how ignoring adjustment costs can attenuate microeconomic elasticities of labor supply (Chetty et al. 2011).

previous case shown in the top panel. These graphs are hypothetical, but illustrative of the ways that budget constraints might influence plan choices.

I now turn to an empirical analysis that compares observed plan choices and HSA saving between employees who took out a 401(k) loans and other employees who did not. One-third of employees take at least one loan from their 401(k) between 2008 and 2009, with 14 percent taking one loan and 18 percent taking two loans. Employees with family coverage are more likely to take a loan than employees with self coverage. Similar to (Lu and Mitchell, 2010), higher-salaried employees are less likely to take loans, but have higher loan balances conditional on taking one. Table 1.11 shows that employees who take 401(k) loans during this period choose lower deductibles, make lower HSA contributions, and withdraw a larger share of HSA assets than employees without loans. Panel A shows the percentage of employees choosing the lowest deductible is over 5 percentage points higher among those with loans than those without loans. This share is higher among employees with self coverage. The highest deductible is chosen less often among those with loans compared to those without loans. These plan choices are consistent with the patterns illustrated in Figure 4.1(b), in which those facing tighter liquidity constraints choose lower deductibles than they otherwise would. Panel B of Table 1.10 shows that the means of deductibles, saving, and salaries are all lower among employees with loans, and the differences are statistically significant. These comparisons provide evidence that plan choices and saving rates are likely influenced by liquidity constraints, at least partly. Yet several pieces of the data also suggests this mechanism plays a small role: (1) the same general patterns in plan choices are observed for those without loans, as the lowest deductible is most popular followed by the highest deductible; (2) the magnitude of the differences in HSA contributions are fairly modest: \$164 for employee contributions and just \$65 for total contributions; (3) HSA disbursements are still large among those without loans, with 67 percent of assets withdrawn annually.

1.8.3. Discounting, mental accounting, and financial literacy

A natural explanation for HSA saving patterns may be discounting. If people have strong preferences for current consumption, they may choose to contribute and withdraw HSA funds to finance current health care costs even though they recognize their future consumption demands. Models of present-biased preferences (Laibson 1997; O'Donoghue and Rabin 1999) help to explain the demand for commitment devices like 401(k)s for retirement saving (Laibson et al. 1998). And yet even the 401(k)—clearly intended by policy as a retirement savings vehicle—still gives account holders ample flexibility for penalty-free withdrawals, known as "leakage" (Beshears et al. 2015; Munnell and Webb 2015). The HSA provides even greater opportunity to draw down assets while working since most consumers incur positive health expenditures each year and have a debit card for withdrawals.

Another possible psychological construct explaining saving and plan choices is mental accounting, which assumes households group income and expenditure items into separate accounts (Thaler 1985, 1990; Prelec and Loewenstein 1998; Shefrin and Thaler 1988). Households maintain a system of category budgeting similar to how businesses have different accounting units. Mental accounting, which is an example of framing, assumes that the marginal propensity to consume differs between various income accounts (e.g. current income, future income, wealth), and that households may also earmark funds for different purposes. The "envelope method" of budgeting—whereby households allocate cash to different physical envelopes for monthly spending on food, gas, etc.—is an example of such behavior. Such behavior is sub-optimal, since if prices change, the most preferred consumption bundle may also change. Holding separate accounts violates the fungibility of money. Prior research has documented violations of fungibility for particular expenditure items like gasoline (Hastings and Shapiro 2013), grocery purchases (Milkman and Beshears 2009), restaurant meals (Abeler and Marklein 2013), and children's clothing (Kooreman 2000). There is also evidence that financial borrowing decisions by some consumers violate the no-arbitrage condition: many take payday loans when lower interest credit is available (Agarwal, Skiba, and Tobacman 2009) or simultaneously hold both high-interest credit card debt and low-yield assets (Gross and Souleles 2002). In the context of HSAs, people may view their HSAs as accounts designated to cover only health care expenses, while their 401(k)s are designed for retirement saving, even though the money is fungible.

Such empirical anomalies may ultimately stem from a gap in financial literacy (Hastings, Madrian, and Skimmyhorn 2013; Lusardi and Mitchell 2014). Financial literacy includes not only information about financial products but also the mathematical skills and conceptual knowledge, such as compound interest, required to make sophisticated financial decisions. In the case of HDHPs and HSAs, consumers must be knowledgable about both the features of health insurance contracts and tax-preferred savings vehicles to make informed decisions. Acquiring such knowledge represents a tall order for many people. Research in other employment settings suggests many HDHP policyholders lack knowledge about what their plan covers (Lieu et al. 2010; Reed et al. 2012) and such information gaps meaningfully impact plan choices (Handel and Kolstad 2015). Employees in this setting appear knowledgeable about HDHP coverage, the tax benefits of HSA saving, and the ability for HSA funds to roll over; however, it is possible they lack knowledge about the fungibility between HSA and 401(k) accounts, which is arguably the most complex feature of these contracts.

1.9. Welfare Implications

1.9.1. Changes in health care consumption and moral hazard

The policy objective when introducing HDHPs was for consumers to face the full cost of care for small expenses, rather than pay a fraction of the total cost with coinsurance. Yet as we have seen, a positive MPC leads consumers to view their HSA saving as a price reduction, undermining the cost-control incentives of HDHPs. This section seeks to approximate the reduction in moral hazard in moving from a traditional plan with a 20 percent coinsurance rate—a benchmark corresponding to the status quo—to an HDHP alone or to an HDHP with HSA saving. The ideal way to measure moral hazard empirically would be to observe whether the quantity of health care consumed differs if insurance provides a lump-sum cash payment or covers the full cost of care. No research has ever run that experiment. Instead, my approach to measuring moral hazard is based on the extent to which consumers choose HSA saving as though the account were cash (i.e. like a 401(k)). The choice of HSA saving is thus similar to plan selection based on anticipated responses to the out-of-pocket price, known as "selection on moral hazard" (Einav et al. 2013). This approach assumes that consumers recognize the fungibility of HSA assets, but as discussed earlier in Section 1.8, the data do not allow me to test this, unlike other potential information frictions.[39]

[39] Another approach to studying moral hazard and HSAs would be to examine whether consumption decisions differ based on the structure of the insurance contract. Suppose a person with a \$500 deductible in traditional insurance chooses to purchase an MRI costing \$1,500, so the person paid \$500 and the insurance plan paid costs of \$1,000. Now suppose that person instead is enrolled in a \$1,500 HDHP and had \$1,000 in their HSA—requiring an additional \$500 for the MRI (ignore the HSA's tax benefits for now). If the person uses the \$1,000 HSA money to buy the MRI, it suggests there is no moral hazard. If the person with the HSA does not buy the MRI, it implies the \$1,000 of benefits financed from traditional insurance were worth less than \$1,000 to the person. I plan to pursue this analysis in future work using data from before and after HDHPs were introduced in this setting. The challenge is to find such clean comparisons empirically. The intuition here is similar to measuring how unemployment durations respond to unemployment benefits for households that are

Let $\alpha \in (0,1)$ denote the coinsurance rate—the fraction of the cost paid by the consumer at the time of service. Figure 6 plots a simple example of the demand for health care, in which the full price of care is denoted by P^* and the price faced by the consumer with coinsurance is αP^*. Assuming the demand curve reflects the marginal benefit of care, there is a welfare loss of area *ABC* from the consumption of care worth less than its cost of production. Moving from insurance with a coinsurance rate of α to an HDHP where $\alpha = 1$ reduces the quantity of care consumed from *Q'* to *Q* and eliminates the welfare loss.

However, a positive MPC leads to smaller decreases in consumption and moral hazard than these benchmark reductions. A positive MPC acts as a price reduction, lowering the full price from P^* to $P^*(1-\eta)$. At this price, the quantity of care consumed is *Q"*, which exceeds the benchmark quantity *Q* when consumers face the full cost. There is an associated welfare loss from moral hazard equal to area *ADE* at this level of consumption. As long as $1 - \eta > \alpha$, HDHPs reduce moral hazard relative to traditional insurance, but the decrease is lower than the benchmark case where HDHPs fully eliminate moral hazard. The reduction in moral hazard when using the HSA to finance current costs is equal to the trapezoid *BCED* in Figure 6 rather than the full welfare loss *ABC*. Following other research, these calculations assume consumers perceive the coinsurance rate as the price of care.[40]

The economic magnitudes of consumption changes will depend on the price elasticity

liquidity-constrained or not to isolate moral hazard separately from liquidity effects (Chetty 2008). Of course, for such behavior to reflect moral hazard, people must understand that HSAs can be used as future cash.

[40] The non-linear nature of the insurance contract implies the shadow price of care will depend on the level of consumption. Rational consumers would consider the expected price paid at the end of the year, which would be zero for those who have reached their out-of-pocket maximum. Recent studies document consumers respond instead to the spot price of care based on where they currently are in the contract's benefit schedule (Abaluck et al.; Dalton et al.; Brot-Goldberg et al.). In my sample, 70 percent do not exceed the deductible, suggesting it is reasonable to consider the marginal price of care equal to one dollar for most people.

of demand and the previous price of care. If we suppose the price elasticity of demand is equal to -0.2, as in the RAND Health Insurance Experiment (Manning et al. 1987; Keeler and Rolph 1988), then moving consumers from a coinsurance rate of 20 percent to paying the full cost under an HDHP would translate into a 27 percent reduction in the quantity of care consumed.[41,42] By contrast, with an MPC equal to 0.68 the reduction in health care consumption is 6 percent.[43]

The reduction in moral hazard when using the HSA to pay for current costs is less than 30 percent as large as the reduction of moral hazard from an HDHP where consumers perceive the full price of care. Applying a variation of the Harberger deadweight loss formula also implemented by Pauly (1969) and Feldstein (1973), the moral hazard cost at coinsurance rate α is approximated as:

$$DWL_{\alpha} = \frac{1}{2}\varepsilon P^{*}Q\left(1-\alpha\right)^{2} \tag{1.16}$$

where ε is the price elasticity of demand. Assuming the same elasticity, the moral hazard cost at the price $P^{*}\left(1-\eta\right)$ is:

$$DWL_{1-\eta} = \frac{1}{2}\varepsilon P^{*}Q\eta^{2} \tag{1.17}$$

The percentage reduction in moral hazard from HSA saving compared to an HDHP alone is therefore approximately equal to $1-\frac{DWL_{1-\eta}}{DWL_{\alpha}} = 1-\frac{\eta^{2}}{(1-\alpha)^{2}} = 1-\left(\frac{0.68}{0.80}\right)^{2} = 0.28$.

[41] This calculation abstracts from the issues involved with choosing a single price when confronted with a non-linear contract (Aaron-Dine et al. 2013). Both for simplicity and comparison to previous work (e.g. Cogan et al. 2011), I choose the coinsurance rate as the price faced by the consumer.

[42] This calculation uses the average price of care for the percentage increase in price (arc elasticity), following Cogan et al. (2011) and others in this literature. Specifically, the 133 percentage price increase is calculated as $(1-0.2)/\left(0.5\times(1+0.2)\right) = 1.33$. Multiplying 1.33 by -0.2 yields -0.266.

[43] This is calculated by again applying the arc elasticity formula to the difference in the price. Specifically, $-0.061 = -0.2\times\left\{(1-0.2)/\left(0.5\times(1+0.2)\right)-(1-0.32)/\left(0.5\times(1+0.32)\right)\right\}$.

The use of HSA saving to pay for current health care costs makes the HDHP 28 percent as effective at reducing moral hazard compared to an HDHP alone.[44] In dollar terms, the HDHP with HSA reduces roughly \$125 of moral hazard compared to a 20 percent coinsurance rate, while an HDHP alone would reduce \$450 in moral hazard. Figure 6 plots this calculation for an MPC ranging from 0 to 1, to depict the percentage reduction in moral hazard relative to a standard 20 percent coinsurance rate.

1.9.2. Consumption smoothing benefits and tax expenditure of HSA contributions

The risk smoothing benefits from using HSAs for additional insurance against the deductible are small relative to the tax expenditure on HSA contributions. Using the choice model's estimates, the risk smoothing benefits are calculated as the difference in certainty equivalents between the chosen deductible and HSA contribution versus the same deductible without an HSA. The mean risk smoothing benefit is estimated to be \$387, which is smaller than the mean tax expenditure of \$591 on HSA contributions. Moreover, this calculation does not incorporate the marginal cost of public funds, which would further increase the cost of taxation.

1.9.3. Impacts on plan choices

Using the HSA to pay for current costs can induce different HDHP choices by allowing people to use HSA saving to reduce the deductible. Table 1.12 presents the results of simulating plan choices using the choice model parameter estimates from Table 9 and compares these choices to outcomes assuming consumers do not have an HSA (so that the MPC is zero by construction). The effect of such plan choices on premi-

44

If demand becomes less elastic as quantity increases (e.g. if the demand curve were linear), this calculation will overstate the percentage reduction in moral hazard.

ums depends on the incremental costs of the risks choosing each plan. As in many settings, the employer contributes a flat amount to the employee's premium based on a percentage of the cost of the highest deductible and sets premiums based on the average expected total cost among consumers enrolled in each plan. Differences in the employee premium between plans is then due to differences in the incremental cost of more generous coverage relative to the lowest highest deductible plan. Without an HSA, the average costs of employees choosing the lowest deductible is predicted to be \$3,283 higher than those choosing the highest deductible for self coverage and \$6,938 higher for family coverage. These incremental differences are smaller in the case with HSAs, in which the differences are \$2,107 and \$5,546, respectively. Since the employee pays for incremental costs relative to the highest deductible, premiums for the lowest deductible would be thus \$1,176 less for self coverage and \$1,392 for family coverage with HSAs compared to without. The differences in incremental costs are also over \$1,000 for the intermediate deductibles in most cases. These patterns suggest that HSAs, at least in this context, may attenuate any adverse selection in insurance choices.[45] These results are highly context-specific and depend on both the schedule of premium benefits as well as the preference parameters of this population. Yet the possibility that HSAs influence premiums by inducing consumer re-sorting across plans is a potentially important and unintended consequence of these contracts.

1.10. Conclusion

This paper studies the experiences of employees using HSAs at a large firm that fully replaced its traditional health insurance offerings with HDHPs. The policy

[45] A complete treatment of selection would entail estimating the slope of the marginal cost curves (Einav et al. 2010). In future work, I intend to study how HSAs influence selection into insurance in greater detail, including the case when consumers can choose between traditional insurance and HDHPs.

rationale behind HDHPs and HSAs was to make consumers more cost-conscious, while offering them a subsidy to finance any form of consumption provided they are willing to wait until age 65. HSAs were designed to reduce costs through tax-preferred saving accounts, but the evidence shows that consumers in this setting use their HSA to reduce their deductible, instead of saving for future consumption. At the margin, 68 cents of every HSA dollar is allocated toward reducing the deductible. Such responses to the tax incentives counteract the cost-control incentives of the high deductible. In this setting, replace traditional coverage with HDHPs produced no decline in health care spending as employees offset the higher deductibles with HSA contributions. Using the HSA to finance short-term health care costs also induces consumers to choose different plans than they otherwise would. Such re-sorting across plans indirectly affects premiums by changing the health risks across plans, which represents an important unintended consequence of HSAs. I find that HSAs mitigate adverse selection in this setting, although the direction and magnitude of any selection will be context-specific based on the population studied and menus of deductibles and premiums.

Consumers face a challenging decision-making problem in response to the contract's tax incentives. HSAs are complex financial accounts that bridge health insurance and retirement saving, which are already extremely difficult decisions. Not only do consumers need financial literacy to make informed insurance and saving choices, but they must also be patient and recognize the substitutability between health care and other goods from HSA funds. In principle, consumer learning about the HSA and HDHPs could lead people to use HSAs as a long-term savings vehicle. An important qualification about this paper's results is that the analysis is limited to only the first two years after HSA adoption. In future work, I plan to study the extent to which consumers learn over time about HSAs.

In conclusion, contracts that require sophisticated consumer decision-making may work well in theory, but may be less effective and lead to unintended consequences in practice. Future research should consider how to design health insurance contracts that achieve their objectives when consumers possibly make optimization errors.

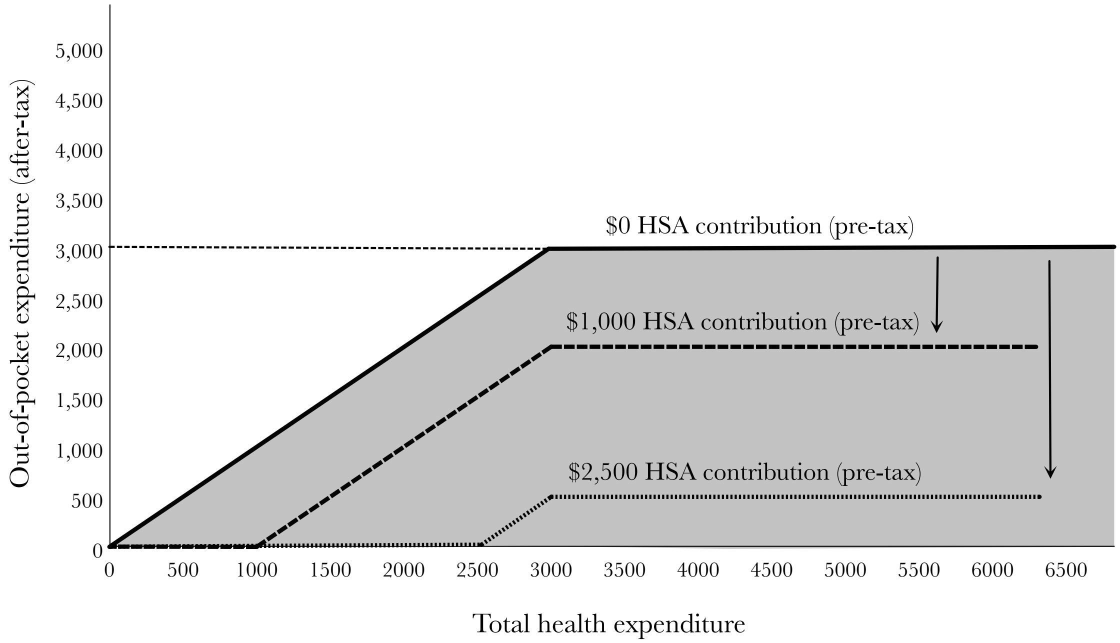
1.11. Tables & Figures

Table 1: Comparison between 401(k) and HSA features

	401(k)	HSA
Contributions exempt from income tax	x	x
Contributions exempt from FICA taxes		x
Interest grows tax-deferred	x	x
Medical care tax free		x
Penalty for early withdrawal before 2011	10%	10%
Age when can withdraw penalty-free	$59\frac{1}{2}$	65
Annual contribution limit (incl. employer)	\$53,000 self	\$3,350 self /\$6,650 family

Note: This table presents the major similarities and differences between HSAs and 401(k)s. In 2011, the penalty for early withdrawn from the HSA was raised from 10% to 20%. The IRS rules for catch-up contributions also differ between accounts. People can contribute an extra \$6,000 starting at age 50 to their 401(k) versus \$1,000 starting at age 55 to their HSA.

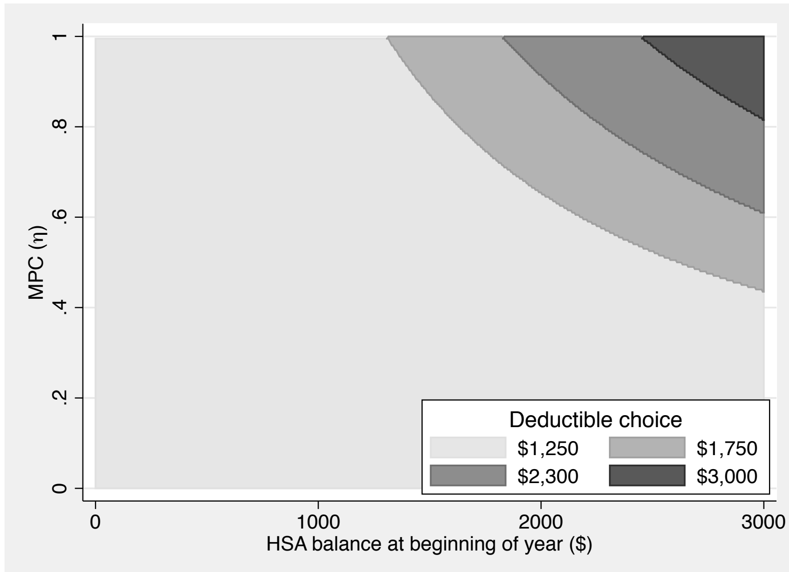
Figure 1: HSA saving for short-term health care costs: continuous choice of insurance contract



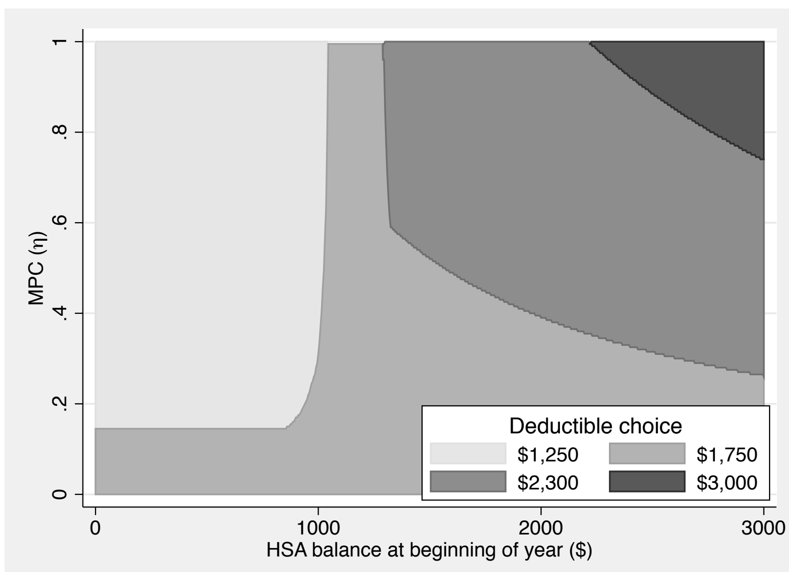
Note: The solid line represents the benefit schedule for a contract with a $3,000 deductible and full coverage past the deductible: the consumer pays for all expenses out-of-pocket until total costs reach $3,000 and the insurance plan pays for all costs beyond this amount. The dashed and dotted lines represent HSA contributions used to provide coverage against out-of-pocket payments associated with the deductible. For example, a $2,500 contribution shifts the benefit schedule towards the horizontal axis, so that the first $2,500 of costs are financed from the HSA contribution with pre-tax funds, the next $500 of costs are financed with after-tax funds, and all costs beyond $3,000 are paid by the insurance plan. With HSA contributions, any contract in the shaded area is feasible so that the choice of contract becomes continuous, in effect.

Figure 2: Example: Deductible choices by MPC, HSA balances, and health risk

(a) Female, age 35-44, Highest Quintile of Health Risk

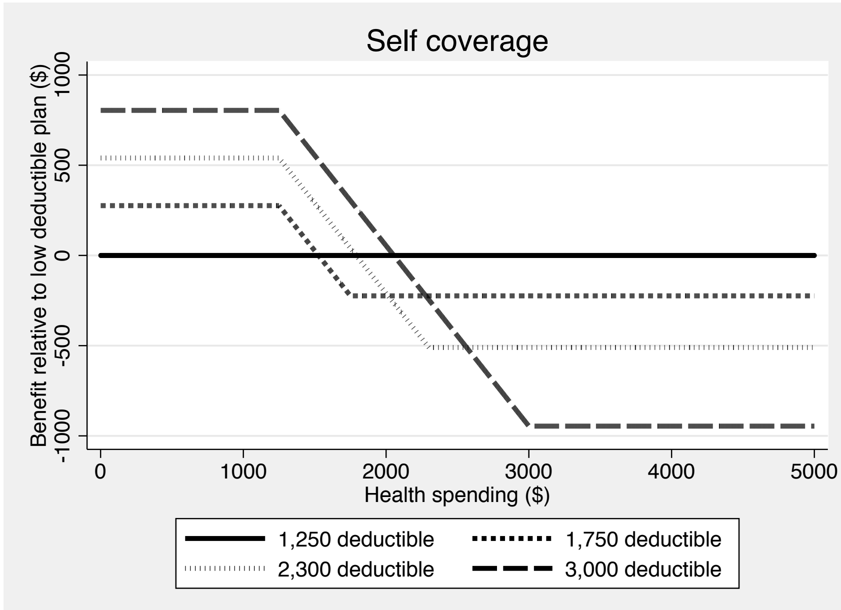


(b) Female, age 35-44, Median Quintile of Health Risk



Note: These graphs plot the chosen deductible assuming a coefficient of absolute risk aversion of $2.5\text{x}10^{-4}$. Panel A uses the distribution of spending for a 35-44 year old female in the highest quintile of health risk and Panel B uses the distribution of spending for a 35-44 year old female in the median quintile of health risk. In Panel A, higher levels of the MPC induce high-cost consumers to choose higher deductibles as HSA balances increase. In Panel B, a positive MPC can induce consumers with moderate health spending to choose lower deductibles if HSA balances are low, or higher deductibles if HSA balances are high.

Figure 3: Benefits relative to low deductible plan by health spending, 2008



Note: Benefits are calculated as health spending less premiums and out-of-pocket payments. Out-of-pocket payments equal health spending below the deductible and zero after the deductible. This figure presents the benefit schedule for self-coverage. Coverage for employee plus spouse, employee plus children, and families exhibit similar patterns.

Table 2: Employer Matching Rates for 401(k) and HSA, 2007-2009

A. 401(k)

Year	Salary level	Match rate m_{401k} (share matched)	Match limit
2008, 2009	All employees	1 on 1% of salary then 0.5 after	6.5% salary

B. HSA

Year	Salary level	Match rate m_{HSA} (share matched)	Match limit (self / family)
2008	$< \$50,000$	6	\$600 / \$1,200
	\$50,000 - \$99,999	4	\$400 / \$800
	\$100,000 - \$149,999	2	\$200 / \$400
	$\geq$\$150,000	0	\$0 / \$0
2009	$< \$50,000$	6	\$600 / \$1,200
	\$50,000 - \$99,999	4	\$400 / \$800
	\$100,000 - \$149,999	0	\$0 / \$0
	$\geq$\$150,000	0	\$0 / \$0

Note: This table presents the employer's matching schedule for employee 401(k) and HSA contributions/ The match rate is defined as the share of the employee's contribution that is matched by the employer. A match rate of 2, for example, denotes that the employer contributes \$2 for each \$1 contributed by the employee. All employees received the same match rate on 401(k) contributions up to 6.5% of salary. By contrast, employees with lower salaries receive a higher match rate and match limit on HSA contributions than employees with higher salaries.

Table 3: Summary Statistics of Sample

	All employees mean	All employees s.d.	Self coverage mean	Family coverage mean
HSA employee contribution ($)	1,141.0	1,095.9	766.9	1,475.4
HSA employer contribution ($)	735.7	349.6	512.9	934.9
HSA balance ($)	318.9	661.6	262.5	369.3
HSA withdrawal ($)	1,495.0	1,098.4	953.9	1,978.7
401(k) employee contribution ($)	3,271.9	3,639.6	2,936.5	3,571.7
401(k) employer contribution ($)	1,556.9	1,366.4	1,388.6	1,707.3
401(k) balance ($)	29,045.8	64,197.0	22,290.8	35,077.7
Deductible ($)	3,068.4	1,520.8	2,028.9	3,997.6
Expected health spending ($)	6,302.0	5,669.4	3,961.1	8,394.6
Salary ($)	54,517.9	34,211.4	49,053.7	59,402.4
Tenure (years)	6.4	5.6	5.97	6.77
Age (years)	39.45	9.9	38.33	40.45
% female	71.3	0.5	73.3	69.5
% married	51.2	0.5	29.8	70.2
% White	66.6	0.5	65.7	67.5
Number of dependents	1.09	1.3	0.0	2.1
% executive/upper	6.0	0.2	3.3	8.5
% manager	29.2	0.5	28.6	29.8
% support staff	28.5	0.5	29.8	27.4
% technician	36.2	0.5	38.3	34.4
% zero HSA contribution	0.8	0.1	0.9	0.7
% obtaining full employer HSA match	96.6	0.2	96.7	96.5
% contributing HSA maximum	3.0	0.2	3.4	2.6
% zero 401(k) contribution	5.7	0.2	5.7	5.8
% obtaining full employer 401(k) match	43.2	0.5	44.8	41.8
% contributing 401(k) maximum	1.9	0.1	0.9	2.7

Note: This table presents means and standard deviations of the analysis sample by type of coverage. Family coverage also includes coverage for employee plus spouse and employee plus children. 1 percent of households switch between self coverage and family coverage over the two-year period. The statistics on HSA balances presented in this table are averaged over 2008 and 2009, although the beginning-year HSA balance was zero for all observations in 2008 because that was the year employees established their accounts.

Table 4: Premiums, Deductibles, and Enrollment for Health Plans, 2008-2009

Self coverage			Family coverage		
Deductible choice ($)	Premiums ($)	Enrollment (% total)	Deductible choice ($)	Premiums ($)	Enrollment (% total)
A. 2008					
HDHP 1,250	1,504	43.7	HDHP 2,500	3,580	42.8
HDHP 1,750	1,228	12.6	HDHP 3,500	2,728	12.9
HDHP 2,300	964	12.4	HDHP 4,600	1,912	17.4
HDHP 3,000	700	31.4	HDHP 6,000	1,084	26.8
B. 2009					
HDHP 1,350	1,669	45.1	HDHP 2,700	3,911	44.5
HDHP 1,850	1,376	11.5	HDHP 3,700	2,975	14.4
HDHP 2,400	1,102	10.6	HDHP 4,800	2,097	12.5
HDHP 3,000	836	32.9	HDHP 6,000	1,246	28.6

Note: This table displays the premiums for each deductible offered in 2008 and 2009 and corresponding enrollment. For visual clarity, premiums for employee plus spouse coverage and employee plus children coverage are not presented, but are included in all analyses.

Table 5: Fixed effects regressions of employee saving

	Dependent variable: Employee HSA saving			Dependent variable: Employee 401(k) saving	
	Levels	Levels	Log-Log	Levels	Log-Log
	(1)	(2)	(3)	(4)	(5)
HSA balance at beginning of year	-0.307***		-0.055***	0.200***	0.089***
	(-12.01)		(-3.68)	(4.09)	(5.12)
Deductible	0.038		-0.159*	0.001	-0.035
	(1.84)		(-2.06)	(0.02)	(-0.35)
Expected OOP less HSA balance		0.213***			
		(10.39)			
Employee HSA saving				0.073	0.085**
				(1.39)	(2.60)
Employee 401(k) saving	0.022	0.012	0.038**		
	(1.37)	(0.64)	(2.63)		
401(k) balance at beginning of year	-0.001	-0.001	-0.024*	-0.002	0.017
	(-1.19)	(-0.37)	(-2.23)	(-1.67)	(1.33)
Price per HSA dollar, $\frac{1}{1+m_{HSA}}$	-393.35	-611.73	0.18	2209.24**	-0.59
	(-1.23)	(-1.90)	(0.96)	(2.51)	(-1.51)
HSA match limit	-2.899*	-3.622*	0.035	8.483**	-0.147
	(-2.09)	(-2.57)	(0.74)	(2.52)	(-1.76)
401(k) match limit	-0.017	-0.011	0.299	0.497**	0.765***
	(-0.62)	(-0.37)	(1.54)	(2.69)	(4.20)
Employee fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
N	8,873	8,873	8,864	8,873	8,864
R-squared	0.131	0.090	0.041	0.109	0.054

Note: All regressions also include expected health spending, cubics in age, salary, and tenure, the number of dependents, family coverage indicator, and a constant. Robust t-statistics clustered by employee in parentheses.
*p<0.05, **p<0.01, ***p<0.001.

Table 6: Plan choice regressions

	Fixed effects logit: Dep. variable = 1 if given deductible chosen, 0 otherwise		Ordered logit: Dep. variable equals deductible chosen	
	Odds ratio		Coefficient estimate	
	Lowest deductible, all coverage	Highest deductible, all coverage	Self coverage	Family coverage
Variables in $100s	(1)	(2)	(3)	(4)
HSA balance at beginning of year	0.969*	1.037**	0.024***	0.016**
	(-2.36)	(2.78)	(3.15)	(3.31)
HSA saving (Employee + Employer)	0.982	1.013	-0.030***	-0.003
	(-1.86)	(1.31)	(-3.79)	(-0.69)
401(k) balance at beginning of year	1.000	1.000	0.000	0.000*
	(0.60)	(-0.90)	(0.32)	(2.12)
401(k) saving (Employee + Employer)	1.009	0.996	0.002	0.001
	(1.60)	(-0.89)	(1.07)	(0.79)
Expected health spending	1.008***	0.992***	-0.027***	-0.013***
	(4.03)	(-3.03)	(-13.26)	(-15.53)
Employee fixed effects	Yes	Yes	No	No
Year fixed effects	Yes	Yes	Yes	Yes
N	1,252	1,088	2,248	2,532
Log-likelihood	-430.3	-383.1	-2,415.9	-2,940.2

Note: Columns 1 and 2 present the results of fixed effects logits of choosing either for the lowest deductible (Column 1) or the highest deductible (Column2). These regressions use within-employee variation and only include employees who switch to or from the lowest or highest deductible choices. A hazard ratio lower than 1 indicates a lower propensity to choose either the lowest deductible or highest deductible, while a hazard ratio exceeding 1indicates a greater propensity. Columns 3 and 4 present ordered logits that use between-employee variation in deductible, HSA balances, and other observables. Regressions also include cubics in age, salary, and tenure. Ordered logits include indicators for white and marital status. Both sets of regressions indicate that higher HSA balances are positively correlated with higher deductible choices. Robust t-statistics clustered by employee in parentheses.*p<0.05, **p<0.01, ***p<0.001.

Table 7: Predicted Probabilities of Deductible Choice from Ordered Logit

A. Self coverage

If HSA balance at beginning of year equals	Probability chosen deductible equals: $1,350	$1,850	$2,400	$3,000
$0	49.4	14.2	11.6	24.8
	[46.3, 52.4]	[12.6, 15.8]	[10.2, 13.1]	[22.4, 27.2]
Median ($288)	47.7	14.4	11.9	26.0
	[45.1, 50.3]	[12.7, 16.0]	[10.5, 13.3]	[23.9, 28.3]
Mean ($476)	46.4	14.4	12.1	27.0
	[44.0, 48.9]	[12.8, 16.1]	[10.6,13.6]	[24.9, 29.2]
75th percentile ($676)	45.3	14.5	12.3	27.9
	[42.7, 47.9]	[12.8, 16.1]	[10.8, 13.8]	[25.6, 30.2]
95th percentile ($1,587)	39.7	14.4	13.1	32.8
	[34.9, 44.5]	[12.8, 16.1]	[11.5, 14.7]	[28.2, 37.4]

B. Family coverage

If HSA balance at beginning of year equals	Probability chosen deductible equals: $2,700	$3,700	$4,800	$6,000
$0	48.3	16.7	12.7	22.3
	[45.5, 51.0]	[15.1, 18.3]	[11.3, 14.1]	[20.1, 24.5]
Median ($325)	46.9	16.8	13.0	23.2
	[44.6, 49.3]	[15.2, 18.4]	[11.6, 14.3]	[21.3, 25.2]
Mean ($683)	45.5	16.9	13.3	24.3
	[43.3, 47.6]	[15.3, 18.5]	[11.9, 14.6]	[22.4, 26.2]
75th percentile ($990)	44.2	17.0	13.5	25.3
	[41.9, 46.5]	[15.4, 18.6]	[12.1, 14.9]	[23.3, 27.3]
95th percentile ($2,397)	38.6	17.0	14.5	29.9
	[34.2, 43.1]	[15.4, 18.6]	[12.9, 16.2]	[25.9, 33.9]

Note: This table presents the predicted probabilities of choosing each deductible in 2009 for different levels of the beginning-year HSA balance using the coefficient estimates from the ordered logit model presented in Table 1.6. The probabilities are calculated holding all other variables at their mean. 95 percent confidence intervals of the predictions are displayed in parentheses. For both self and family coverage, higher balances predict higher deductible choices, all else equal.

Table 8: Within-Employee Regressions of Log Total Health Care Spending

	All employees	Similar OOP max before and after HDHP	Similar OOP max before and after HDHP	Same OOP max before and after HDHP
	(1)	(2)	(3)	(4)
HDHP and HSA (1=yes, 0=no)	0.060*	0.018	-0.037	-0.053
	(0.032)	(0.074)	(0.078)	(0.095)
Salary measured in:	Deciles	Deciles	Logs	Deciles
Expected health spending measured in:	Deciles	Deciles	Logs	Deciles
Employee fixed effects	Yes	Yes	Yes	Yes
N (household-years)	14,464	4,795	4,795	3,152

Note: Regressions also include number of dependents and a constant. Robust standard errors clustered by employees in parentheses. *p<0.05, **p<0.01, ***p<0.001.

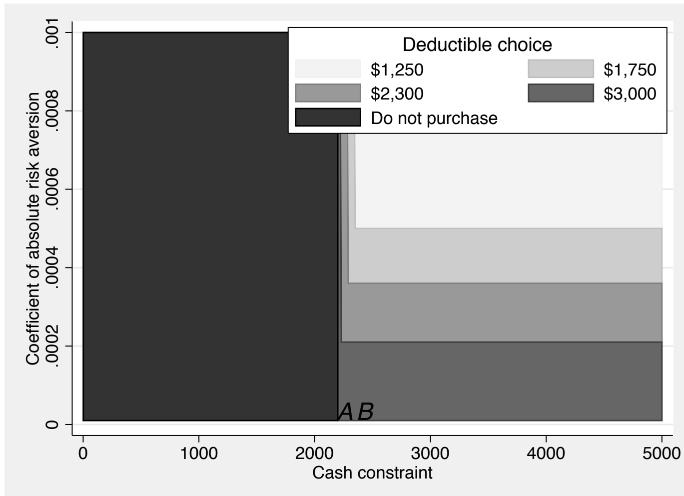
Table 9: Parameter Estimates from Choice Model

	Standard model (1)	HSA model (2)
Risk aversion mean μ_γ	$\mathbf{1.5 \times 10^{-3}}$	$\mathbf{3.5 \times 10^{-4\dagger}}$
Gamble interpretation	**$383**	**$739†**
Risk aversion - intercept α_γ	1.0×10^{-6}	4.8×10^{-9}
Risk aversion - age (years) slope β_γ^{age}	3.8×10^{-6}**	2.5×10^{-6}**
Risk aversion - salary ($1,000s) slope β_γ^{sal}	-3.0×10^{-5}	-2.1×10^{-6}**
Risk aversion - indicator for high tenure β_γ^{tenure}	4.1×10^{-6}**	-1.8×10^{-4}
Risk aversion - indicator for finance job β_γ^{fin}	-4.1×10^{-6}**	-2.5×10^{-5}
Risk aversion - indicator for self coverage β_γ^{self}	2.8×10^{-4}	5.1×10^{-5}**
Risk aversion std. deviation σ_γ	3.6×10^{-3}	7.6×10^{-4}
MPC mean η		**0.68****
MPC 95 % CI		**[0.65, 0.72]**
MPC intercept α_η		9.5×10^{-2}
MPC - age (years) slope β_η^{age}		1.4×10^{-2}**
MPC - salary ($1,000s) slope β_η^{sal}		-1.2 $\times 10^{-3}$
MPC - indicator for high tenure β_γ^{tenure}		1.2×10^{-2}
MPC - indicator for finance job β_η^{fin}		8.6×10^{-3}
MPC - indicator for self coverage β_η^{self}		1.6×10^{-1}**
$\sigma_{\epsilon 1750}$ self-only coverage	54.57	127.32
$\sigma_{\epsilon 2300}$ self-only coverage	10.35	151.73
$\sigma_{\epsilon 3000}$ self-only coverage	527.03	6.64
$\sigma_{\epsilon 1750}$ family coverage	141.89	218.65
$\sigma_{\epsilon 2300}$ family coverage	42.18	78.99
$\sigma_{\epsilon 3000}$ family coverage	734.97	10.52
401(k) saving	No	Yes
Likelihood ratio test stat. versus (1)		310.2

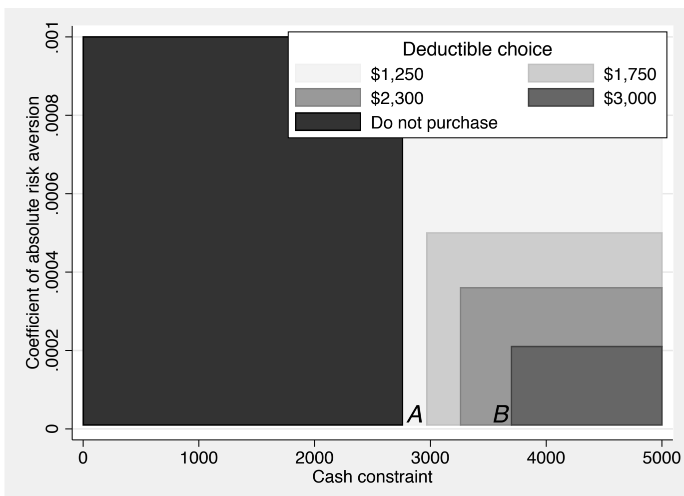
Note: This table presents results from the choice model estimated via simulated maximum likelihood. Column 1 presents estimates from the standard model that does not include HSA saving. Column 2 presents estimates from the model that includes HSA and 401(k) saving to estimate the MPC in addition to risk aversion. In both columns, the row labeled Gamble Interpretation displays the amount Y such that a consumer with the given level of risk aversion would be indifferent between accepting and rejecting a 50-50 gamble of winning $1,000 and losing Y. Standard errors calculated by bootstrapping. ** denotes 95 percent confidence interval excludes zero.
† denotes point estimate is statistically different from same parameter in column 1 at 5 percent level.

Figure 4: Deductible choices by risk aversion and liquidity constraints, self coverage in 2008

(a) Cash constraint: Premiums + Expected Out-of-Pocket Payments

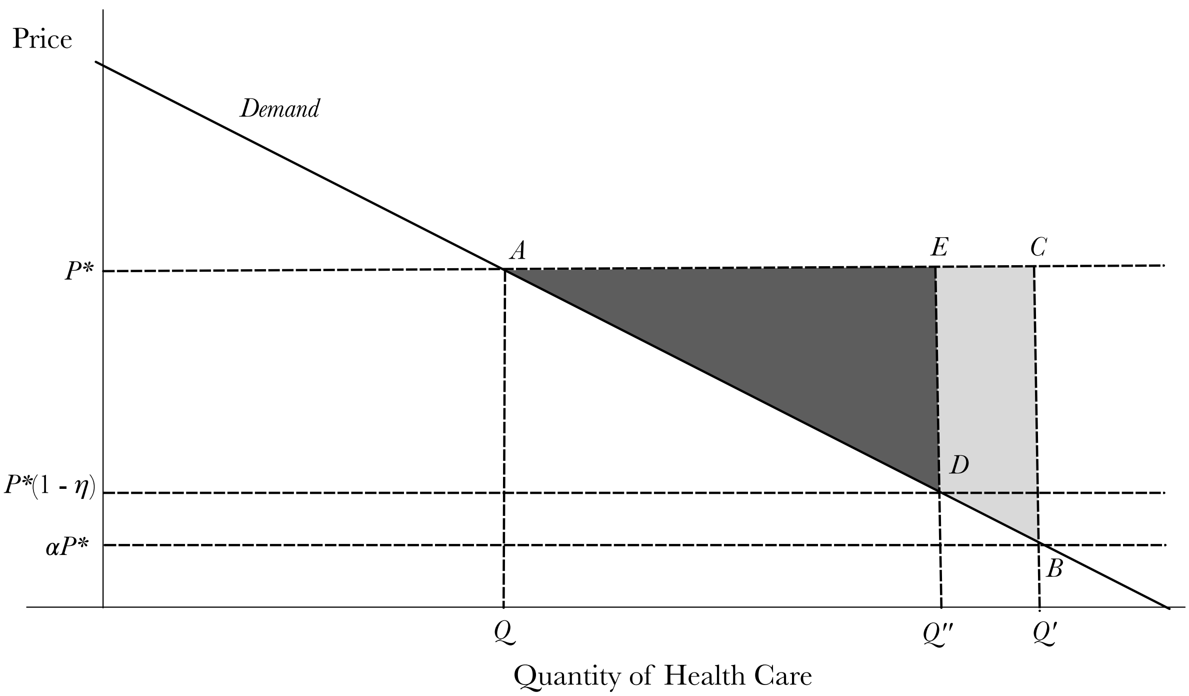


(b) Cash constraint: Premiums + Deductible



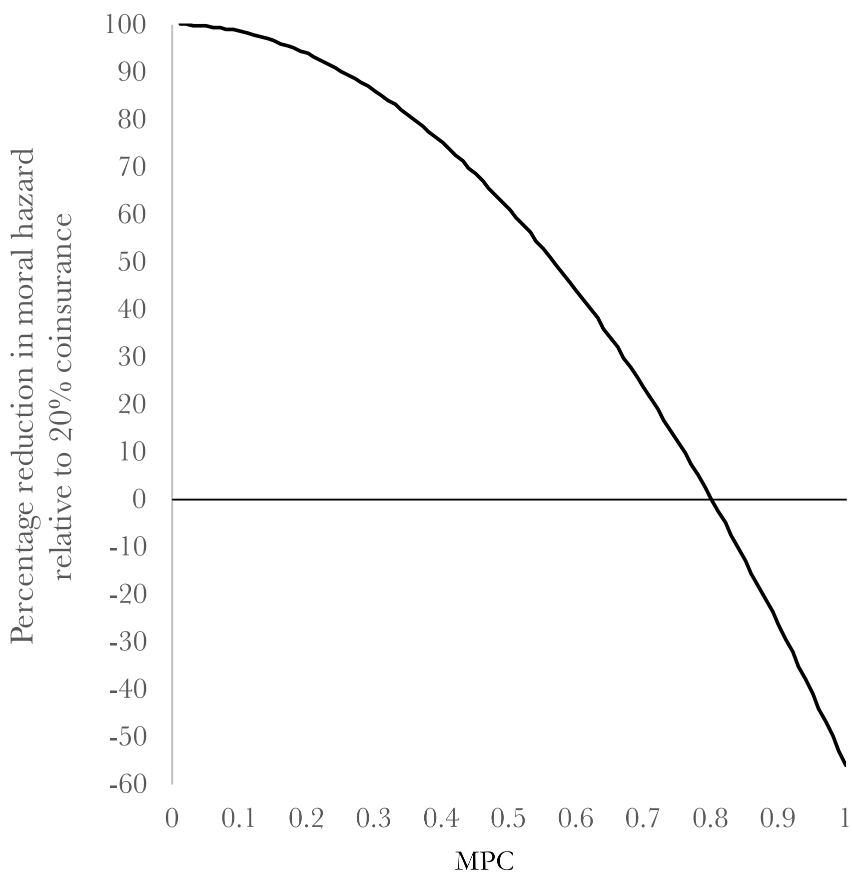
Note: These figures display predicted deductible choices for different levels of risk aversion and the cash constraint for health care using the cost distribution of a female aged 35-44 with the median level of health care costs. Figure (a) defines cash constraints as the sum of premiums and expected out-of-pocket payments. Figure (b) defines cash constraints as the sum of premiums and the deductible. In both graphs, the points A and B denote the range over which cash constraints may distort plan choices. If the budget for health care is below point A, no insurance plan can be purchased. If the budget for health care is larger than point B, plan choices only differ by risk preferences. Between A and B, the employee may not be able to purchase their preferred plan due to their cash constraint. This window is quite narrow (equal to $101) if the cash constraint is defined based on the sum of premiums and expected out-of-pocket payments, since these differ only slightly between plans. The window is wider (equal to $946) if consumers must be prepared to finance the entire deductible, as shown in Figure (b). In that case, consumers may choose lower deductibles not because of risk aversion but because of liquidity constraints.

Figure 5: The quantity of health care demanded when using HSA saving as a price reduction



Note: This figure depicts the change in quantity of health care demanded in response to changes in price. When the the marginal propensity to consume from HSA assets, defined as η, is greater than 0, then HSA saving functions as a price reduction on current health care costs. If the consumer pays the full cost of care at price P^*, in which there is zero moral hazard, then the quantity demanded is Q. A traditional insurance plan charges a coinsurance rate $\alpha \in (0, 1)$, which induces consumption at the level of *Q'*. At this level of consumption, the welfare loss from moral hazard is equal to area *ABC* assuming the demand curve reflects the marginal benefits from care. If an HDHP/HSA contract were to eliminate moral hazard, consumption would decrease from *Q'* to Q. However, using the HSA to reduce current health care costs reduces the price under an HDHP/HSA contract from P^* to $P^*(1 - \eta)$. The reduction in moral hazard is therefore the smaller area *BCED.* In this illustration, the previous coinsurance rate α is assumed to be smaller than η, in line with the model's estimates. This inequality is likely to hold in most settings because a standard coinsurance rate is often 0.1 or 0.2. However, if the MPC were to exceed the previous coinsurance rate, then health care consumption and moral hazard would *increase* in moving from a traditional insurance plan to a HDHP/HSA contract.

Figure 6: Reduction in moral hazard as a function of the MPC



Note: This curve calculates $1 - \frac{DWL_{1-\eta}}{DWL_{\alpha}}$ assuming a coinsurance rate of 20 percent ($\alpha = 0.2$) for MPC ranging from 0 to 1, using the formulas in equations (16) and (17). The negative slope of the curve demonstrates that a higher MPC leads to smaller reductions in moral hazard. In the extreme case that the MPC equals zero, then moral hazard would be eliminated. If the MPC were equal to 0.8, then the consumer perceives the price of care as equivalent to that from a 20 percent coinsurance rate, leaving the level of moral hazard unchanged. If the MPC exceeds $1 - \alpha$, then moral hazard would increase because consumers perceive a lower price of care with an HDHP/HSA than with a 20 percent coinsurance rate.

Table 10: Coefficient of absolute risk aversion from calibration using HSA withdrawals

Quantity	Estimate	Gamble Interpretation
5th percentile	1.0×10^{-6}	\$999
10th percentile	1.0×10^{-6}	\$999
25th percentile	2.0×10^{-6}	\$998
Median	2.4×10^{-4}	\$805
75th percentile	8.0×10^{-4}	\$548
90th percentile	3.2×10^{-3}	\$210
95th percentile	5.0×10^{-3}	\$137

Note: This table presents results from the calibration that point-identifies risk aversion using deductible choices, HSA withdrawals, marginal tax rates, and estimated cost distributions. Unlike the choice model, this exercise does not make a distribution assumption on risk aversion or add i.i.d. plan-specific normal errors. The column labeled Gamble Interpretation displays the amount Y such that a consumer with the given level of risk aversion would be indifferent between accepting and rejecting a 50-50 gamble of winning \$1,000 and losing Y. The median level of risk aversion of 2.4×10^{-4} is close to the estimate from the choice model presented in Column 2 of 8, providing support to the main results. The mean is less informative in this exercise because it is affected by the upper and lower bounds of the grid points chosen for the calibration.

Table 11: HSA saving and plan choices by 401(k) loan status, 2008-2009

Panel A. Plan Choices by 401(k) loan status (%)

	Self coverage		Family coverage	
	With 401(k) loan (N=1,102)	No 401(k) loan (N=3,083)	With 401(k) loan (N=1,534)	No 401(k) loan (N=3,170)
Lowest deductible	49.0	42.8	47.0	42.1
2nd-lowest deductible	12.6	11.7	12.2	14.5
3rd-lowest deductible	10.3	11.8	14.3	15.0
Highest deductible	28.1	33.7	26.5	28.4

Panel B. t-tests of differences in means

	With 401(k) loan, avg.	No 401(k) loan, avg.	p-value from t-test
Deductible, self-coverage	$1,934	$2,067	0.000
Deductible, family coverage	$3,932	$4,033	0.037
HSA saving, employee	$1,022	$1,186	0.000
HSA saving, employee + employer	$1,825	$1,890	0.016
HSA disbursements	$1,602	$1,443	0.000
Share of HSA assets withdrawn	0.79	0.67	0.000
Salary	$50,075	$56,252	0.000

Note: This table presents statistics on plan choices, saving, and salaries for employees who take at least one 401(k) loan in 2008 or 2009 and for other employees without a loan. Panel A lists the market shares by plan for self and family coverage, consolidating choices in years 2008 and 2009 and grouping deductibles by their relative size within each year. Employees with a 401(k) loan choose the lowest deductible plan more often and the highest deductible plan less often. Panel B shows means for deductibles, HSA and 401(k) saving, and salary and reports the p-values of t-tests that the means are equal. In all cases, the means are statistically different, with employees with a 401(k) loan choosing lower deductibles, saving less, and earning less than employees without 401(k) loans.

Table 12: Market shares, average cost, and premiums of insurance plans by MPC

Plan choice, 2009	Market shares (%)		Average total cost ($)		Incremental cost relative to highest deductible ($)	
	$\eta = 0$	$\bar{\eta} = 0.68$	$\eta = 0$	$\bar{\eta} = 0.68$	$\eta = 0$	$\bar{\eta} = 0.68$
			A. Self coverage			
HDHP $1,350	35.9	50.1	5,692	6,080	3,283	2,107
HDHP $1,850	45.1	21.4	6,136	5,939	3,727	1,966
HDHP $2,400	15.1	18.1	4,047	4,332	1,638	359
HDHP $3,000	4.0	10.4	2,409	3,973	-	-
			B. Family coverage			
HDHP $2,700	30.3	46.7	11,261	11,345	6,938	5,546
HDHP $3,700	58.0	37.2	10,439	10,039	6,116	4,240
HDHP $4,800	6.6	8.5	5,601	6,889	1,278	1,090
HDHP $6,000	5.2	7.7	4,323	5,799	-	-

Note: This table calculates the market shares and average total costs without an HSA ($\eta = 0$) and with an HSA with the MPC estimates reported in Table 1.9, which vary by age, income, and job type ($\bar{\eta} = 0.68$). Each row in the table corresponds to a different choice of insurance plan.

CHAPTER 2 : Dying to Win? Olympic Gold Medals and Longevity

2.1. Introduction

Competition for status is ubiquitous in both professional and social settings. This paper studies how status competition affects long-term health. Disentangling the relationship between status and health is challenging because several channels may operate simultaneously. First, higher status can directly expand income opportunities or other real resources that impact health. Second, higher status may produce psychological effects on health, through changes in stress levels, for example. Third, the very pursuit of status may harm health: time spent working may crowd out labor inputs to health like exercise, or conspicuous consumption may displace inputs purchased in the market like medical care. Fourth, obtaining higher status may affect future motivation and thereby influence real resources and health. Finally, a third variable, such as latent ability, could independently determine both status and health.

The existing literature has struggled to separate these mechanisms. The Whitehall studies of British civil servants provide epidemiological evidence of a positive relationship between status and health in an employment setting (Marmot et al., 1991), but endogenous selection into jobs suggests causality does not run from status to health (Chandra and Vogl, 2010; Case and Paxson, 2011). Other research focusing on well-defined occupations in which status is based on receiving awards—Nobel laureates, Oscar winners, and Major League Baseball Hall of Famers—also tends to find a positive association between status and longevity (Sylvestre, Huszti, and Hanley, 2006; Becker, Chay, and Swaminathan, 2007; Rablen and Oswald, 2008). However, unobserved heterogeneity between winners and losers and the non-random assignment of status from these contests raises doubts that the results should be interpreted as measuring the effect of status on health.

In this paper, I compare mortality between Gold and Silver medalists in Olympic Track and Field between 1896 and 1948 to overcome these challenges. While the setting is highly specific, its institutional features provide advantages that help to cleanly identify status and distinguish between the channels listed above. Track and Field includes events in running, jumping, and throwing that use only time or distance to objectively measure performance. In each event, the order of finishers creates a clear and undisputed ranking, even though the differences between competitors may be just fractions of a second. The stakes of such competition are high, with an Olympic victory representing the pinnacle of the sport and carrying global recognition. Variation in status is based simply on winning or losing.

Conditional on reaching the Olympic final, randomness plays a larger role in deciding the difference between winners and losers compared to contests judged over a longer time period. The Olympic Gold medalist is determined on a single day every four years. As I later document, the athlete with the best performance in the year prior to the Olympics often fails to win the Olympic final. Additionally, more than half of athletes who set a World Record never win Olympic Gold.[46] Prior success clearly does not guarantee victory in the Olympics.

Another advantage of this setting is that athletes are physically similar in terms of their baseline health by virtue of their participation in the Olympic final. As supporting evidence of this claim, I show that differences in ability between Olympic finalists—which may positively correlate with both health and winning—do not predict mortality by comparing athletes who ever held World Records (the highest ability

[46] This statistic excludes athletes who competed during the years when the Olympics was canceled due to WWI and WWII and so is not artificially deflated.

group) to those who never did. Since athletes are generally young during Olympic competition, there is also less concern that results are biased by reverse causality in which health determines status. However, performance-enhancing drugs (PEDs) complicate this relationship to the extent that PEDs influence both health and the chance of winning. Since it is difficult to determine which athletes use PEDs, I restrict my analysis to the period 1896 to 1948, when there was less suspicion or evidence of PEDs in Olympic competition.[47]

In addition, income directly earned from competition during this period was non-existent due to the prevailing system of amateurism, which prevented athletes from receiving financial compensation tied to their performance. Until the 1980s, regulations prohibited professional athletes from competing in the Olympics and most Olympians held other occupations while training. The Gold medal itself was also worth a modest amount in terms of its metallic content (Economist, 2012).

Matching data on Olympic finishing order with each athlete's date of birth and death, I first document that Gold medalists die two years earlier than Silver medalists. I estimate survival models that control for observables like height, weight, country, event, and year of birth, which may be correlated with both finishing place and longevity. I then analyze supplementary data on pre-Olympic performances and post-Olympic earnings to test whether this pattern can be explained by an athlete's Olympic performance relative to his expectations—representing psychological factors—or by income—representing real resources.

[47]The International Olympic Commission first produced a list of banned substances in 1968. Some athletes experimented with substances to improve performance that also had health effects in the early 1900s, although doping strategies were not yet advanced. For example, George Hicks won the 1904 marathon after consuming raw egg, Strychnine (a poison that also functioned as a stimulant), and brandy. Drugs yielding significant performance benefits like anabolic, androgenic steroids were not used until the 1950s and amphetamines not until the 1960s (Wadler 1998, WADA 2010).

Comprehensive data on annual rankings by event are used to construct each athlete's expected Olympic finish based on his pre-Olympic performances. An athlete's pre-Olympic rank serves as a clear reference point in this setting given the objective nature of competition. Gold medalists were ranked several places higher than Silver medalists prior to the Olympics. So while Gold medalists by definition either met or exceeded their expectations, Silver medalists outperformed their pre-Olympic rank to a greater extent. Based on these rankings, Gold medalists expected to win while Silver medalists were more often dark horses. The degree to which an athlete outperforms expectations is positively correlated with lifespan, controlling for finishing place. Importantly, the pattern also operates in reverse: Silver medalists who were ranked first before the Olympics died earlier than other Silver medalists who were not considered favorites. These findings are consistent with models of utility based on success relative to a performance benchmark (Rayo and Becker 2007a,b) and expectations-based reference dependence (Koszegi and Rabin, 2006).

While performance relative to expectations is positively correlated with lifespan, this metric adds limited explanatory power and cannot fully reconcile the empirical pattern between winning and a shorter life. In the baseline model, 30 percent of the variation in lifespan is explained by finishing place and other observables. Including the measure of relative performance increases this share to just 35 percent, and the coefficient estimate on finishing place remains large and statistically significant.

There is more empirical support for income as the key mechanism between winning and a shorter lifespan. Focusing on the sub-sample of U.S. athletes, I collect data on earnings and occupational choices for athletes appearing in the 1940 U.S. Census, which was the first Census to record income. Compared to Gold medalists, Silver medalists earned higher incomes and were more likely to enter professional occupations. Losing is again correlated with a lower hazard of death in this sub-sample, but this effect disappears once income is accounted for. There is a large and statistically significant association between higher income and a longer lifespan. Including income explains 60 percent of the variation in lifespan, roughly twice as much as baseline models with finishing place and other observables alone.

Additional analysis of the Census records of each athlete's family provides suggestive evidence that economic conditions in childhood were likely similar between Gold and Silver medalists. I link athletes appearing in the 1940 Census to their family's earlier records in the 1910, 1920, and 1930 Censuses and collect the occupation of each athlete's parents. Using constructed earnings estimates by occupation from the Integrated Public Use Microdata Series (IPUMS), I test whether parental earnings were equal between Gold and Silver medalists and fail to reject the null of no difference in means. In regression models, the coefficient estimate on the athlete's income remain a statistically significant predictor of lifespan while the estimate on parental earnings is not. Based on a range of specifications and within-family income differences, it is unlikely that omitted variable bias explains the relationship between losing and health. The analysis of Census records of each athlete's post-Olympic earnings and his family history is consistent with relative rank influencing motivation. The data does not allow me to distinguish, however, whether losing motivates or winning demotivates.

The estimates are robust to estimating various parametric and semi-parametric survival models that make different assumptions about unobserved heterogeneity. The results are also not sensitive to dropping any single year or country from the sample. The correlation between winning and lifespan is larger in Olympic Games that were more heavily publicized and in premiere events (such as the 100 meter dash), in which any effect of winning on status is arguably greater. Finally, there is no evidence that news coverage, which may correlate with both lifespan and income, predicts lifespan based on analysis of textual data from U.S. newspapers.

This paper's results challenge conventional wisdom and the conclusions from existing studies that being awarded higher status necessarily improves health (Marmot et al. 1978, 1991; Sylvestre, Huszti, and Hanley 2006; Becker, Chay, and Swaminathan 2007; Rablen and Oswald 2008). Instead, losing can have positive, first-order effects on longevity. While Olympic Track and Field is a highly stylized setting, many people face pivotal life events defined by either success or failure. This study may thus have broader implications for understanding how the binary outcomes of important trials in life can produce long-lasting consequences for health.

2.2. Status, Health, and the Olympics

Researchers cannot randomize people into groups of high and low status and measure how long they live. Experimental studies among non-human primates, however, provide some insights into how random changes to status affect health. Some biological research shows that higher status can improve psychological and physical health by reducing stress. One study of rhesus macaques pinpointed the molecular mechanisms behind such psychosocial responses, demonstrating that manipulating social status (dominance rank) affects gene regulation tied to immune defense (Tung et al., 2012). By contrast, other biological studies find that under certain conditions, such as when the hierarchy is unstable, the highest-ranking animals experience the greatest stress from psychosocial factors (Sapolsky, 2005). Yet even experimental studies have not tracked the longevity of animals at different ranks of a randomly assigned hierarchy to study long-term health outcomes.

Observational research in humans—mostly from the epidemiology and medical literature—has generally found a positive gradient between health and status. The Whitehall Study of British Civil Servants in the 1960s and its second iteration in the 1980s demonstrate a marked social gradient in health across different ranks of government employees (Marmot et al., 1978, 1991, 2001; Marmot and Feeney, 1997). Conventional risk factors explain only one third of the difference in mortality risk between clerical and administrative grades (Marmot and Brunner, 2005).[48] The later Whitehall research focuses on social support and the organization of the workplace as possible channels between status and health.

While the Whitehall research clearly reveals an important and sizable link between status and health, the likelihood of endogenous selection into Civil Service ranks raises concerns about how to interpret the results. It is difficult to disentangle the extent to which higher status led to better health or whether better unobserved initial health led to or was otherwise correlated with higher status (Chandra and Vogl, 2010). As evidence of selection, Case and Paxson (2011) find that current self-assessed health in the Whitehall II sample predicts future civil service grade, but current civil service grade does not predict future self-assessed health. In addition, some research also disputes the mechanisms between status and health analyzed in the Whitehall research. A prospective cohort study of Finnish industrial employees found that low predictability at work was highly correlated with heart attack risk, but other organizational factors highlighted by Whitehall—such as low decision autonomy at work—were not (Vaananen et al., 2012).

[48]For heart disease, for example, clerical workers faced a relative risk of dying that was 2.2 times higher than senior administrative staff, and 1.6 times higher than employees in intermediate professional and executive positions.

Outside of Whitehall, research has examined major shocks to status from receiving awards, such as winning the Nobel Prize (Rablen and Oswald, 2008), election to the Major League Baseball (MLB) Hall of Fame (Becker, Chay, and Swaminathan, 2007), or receiving an Oscar (Sylvestre, Huszti, and Hanley, 2006). The assumption behind these studies has tended to be that status should improve health, and there is evidence of this for the Nobel Prize and MLB Hall of Fame but inconclusive results for Oscar winners. However, unobserved heterogeneity and the process of choosing winners may limit what can be drawn from the findings. For example, the physical attributes of Oscar nominees differ in ways that affect their health, and bias may stem from correlation with the likelihood of winning an Oscar. People may also undertake different lifestyle decisions, follow different diets, and value their health in unobserved ways. The same might be said for Nobel laureates and their peers. Moreover, actors, baseball players, and academics are all professionals who can be financially compensated for their work. Higher income associated with status may thus confound comparisons of longevity within these populations. Since Track and Field athletes in the early 1900s were all amateurs, the Olympic setting does not face this problem. Another issue is that these other studies judge performance over a longer time frame. For example, baseball players nominated for the Hall of Fame are assessed over their entire career. The long duration of such assessment increases the chance that the factors that lead people to succeed may be correlated with their mortality prospects. It is reasonable to believe that there is less unobserved heterogeneity among Olympic athletes within any given event than among Nobel laureates, Oscar nominees, or MLB players.[49]

[49]Not surprisingly, the longevity of Olympians is greater than the general population. Clarke et al. (2012) document that Olympic medalists across all sports live almost 3 years longer than other people of the same age, sex, and country. The research did not explore the reasons for this difference, which might be due to genetics, exercise, diet, income, status, or other factors.

Although there is limited economic research on how status affects health, economists have studied the importance of social comparisons both from theoretical and empirical perspectives.[50] Recent models have incorporated peer comparisons and habit formation, rather than assuming utility depends only on absolute consumption levels. In this way, status conveys hedonic value as economic conditions are compared to a benchmark that may depend on personal history, expectations, and the success of one's peers. For example, Rayo and Becker (2007a,b) develop a model in which agents adjust to a time-varying reference point based on habits and peer comparisons, and make output choices given their preferences. Depending on the underlying parameters, agents may fully adjust to their new reference point quickly or they may habituate only partially and gradually over time. The speed of habituation affects how the agent views his current economic conditions compared to past successes or failures.

In the Olympic setting, an athlete who performs better than expected and loses may feel more content than an athlete who expected to win but finishes second. Studies in psychology have examined the facial expressions of Olympic medalists as shown on television to study their reaction soon after the event (Medvec et al., 1995; McGraw et al., 2005). McGraw et al. (2005) focus on the role expectations of finishing place from Sports Illustrated previews and the athlete's performance in the qualifying rounds. Laboratory participants rated the facial expressions of athletes surpassing their expectations as happier than if the athlete fell short of expectations. The study points to the importance of counterfactual comparisons and emotions immediately after competition. Yet performance in qualifying heats is a poor measure of expectations because to win an Olympic final, athletes aim to conserve energy while

[50]See for example Frank (1985); Easterlin (1995); Clark and Oswald (1996); Falk and Knell (2004); Luttmer (2005); Rayo and Becker (2007b); Heffetz and Frank (2011); Heffetz (2011).

advancing through qualifying rounds. My empirical analysis measures expectations based on performances in pre-Olympic competition, which is a better metric of an athlete's beliefs regarding his finishing place in the Olympic final.

Lifestyle decisions and occupational choices represent a potentially important channel between status and health. In analyzing obituaries published in the New York Times, Epstein and Epstein (2013) find that actors, singers, musicians, and athletes die several years earlier than academics, politicians, business executives, and other professionals. The study suggests the earlier death of the former group may result from greater fame and risky behaviors. If success permanently shifts an agent's reference point, risky activities may be rational attempts to attain the utility achieved at the peak of professional success. After Olympic athletes are finished competing, their occupational choices and lifestyle decisions likely affect their health too. Such choices relate to how winning and losing influences motivation. Evidence from field experiments is mixed on whether information about rankings is motivating or discouraging; some research suggests peer comparisons improve future performance (Tran and Zeckhauser 2012), while other studies find informing employees of their relative rank reduces future effort (Barankay 2012a,b).

2.3. Setting and Data

I focus on the setting of Track and Field because it is the oldest sport where performance is objectively measured.[51] Fewer nations and fewer athletes compete in Swimming or Cycling than in Track and Field. I analyze male athletes only since women did not compete in Olympic Track and Field until 1928, with some events

[51] The only sport for the first 13 of the ancient Olympic Games that began in 776 BC was a 1-stadium length sprint—called the "stadion"—measuring 192 meters (Perrottet, 2004).

being limited to males until the 1990s.

The data includes the order of finish in the Olympic final for each event, the country the athlete competed for, and athlete's birth and death dates, collected from a request to the Olympic Studies Center of the International Olympic Committee and the site SportsReference.com. I focus on comparing Gold to Silver medalists in the spirit of a regression discontinuity design. For most athletes, the data also includes height (measured in centimeters) and weight (measured in kilograms) at the time of the Olympic Games. I calculate lifespan as the number of days between the athlete's dates of death and birth. In robustness tests in Section 2.7, I obtain similar results using the date of the Olympics instead of the athlete's date of birth to calculate lifespan.

I classify "high ability" athletes as those who ever held multiple World Records, including after the Olympic Games. Ability may be positively correlated with both winning and latent health. I use two or more World Records as the threshold for high ability in case athletic performances from the tails of a distribution are due to random variation. An advantage of this metric for ability is that it is clearly defined.[52]

I impose several sample restrictions to cleanly focus on the relationship between Olympic performance and lifespan. I exclude athletes with recorded deaths due to war, accidents (car or plane), and murders because these causes are arguably exogenous and unrelated to behavior. I am instead interested in any behavioral effects possibly associated with finishing place, so that deaths may be endogenously determined by the athlete's performance in the Olympics. The site SportsReference.com maintains a list of deaths from such causes that I draw from. Since this list may not

[52]The personal bests of each athlete over their career could be another way to control for ability, but these are highly collinear with year of birth since each event's top performances improve over time.

include all accidental deaths, I further exclude any athlete who died before age 50 to avoid analyzing early deaths. I also drop athletes whose date of death is missing since I am unable to verify their death. Some of these athletes may still be alive, but excluding them is a safer strategy since athletes who win Olympic medals are more likely to have a recorded date of death than other finalists. As a result, no athletes are censored in my data. This reduces my sample size by less than three percent.

Finally, I concentrate analysis on athletes who finish in the top two in a single Olympic Games and single event, which constitutes the majority of Olympians. If athletes compete in multiple Games, it is not clear how an athlete values each performance relative to the other (e.g. whether the best or the final performance is more salient). Since over three-quarters of the sample competes in a single Olympics, focusing on these athletes provides a standard perspective. The results are robust to including athletes who compete in multiple Olympic Games. After these restrictions, the final sample includes 170 athletes with complete dates of birth, death, and finishing place.

Table 13 presents descriptive statistics of the sample, based on whether the athlete won the Gold or Silver medal. Silver medalists live over 2 years longer than Gold medalists, on average. This difference is statistically significant as shown by the final column. Other observable characteristics such as year of birth, height, weight, and ability (as measured by holding multiple World Records) are balanced between winners and losers.

2.4. Methods

The main identification strategy compares longevity between winners and losers in each event and year of the Olympics. The hazard models outlined below include indicators for losing, event, and year to exploit within-event-year variation in lifespan, as well as the athlete's year of birth. Some models also include an indicator for ability, defined as ever holding multiple World Records as described in Section 2.3. This model does not control for country since there are few cases where pairs of winners and losers in the same event and year compete for the same country. As an alternative identification strategy, I compare the longevity of winners and losers within countries and broad event classes, where similar events are grouped into sprints, middle distance, distance, throws, field, or racewalk. I aggregate individual events in this specification since including fixed effects for both countries and events may create an incidental parameters problem in non-linear models.[53] To capture heterogeneity in body types within event classes, I control for height in this specification.[54] These models also condition on ability and year of birth as in the main specification. As shown in Section 2.5, the results are similar under both specifications, but I focus on models using within-event-year identification since that comparison more directly represents the cutoff between winning and losing.

I model lifespan using parametric and semi-parametric hazard models. While running OLS on log life expectancy yields similar results, I present hazard models because the data generating mechanism for mortality is likely to be Gompertz rather than log normal. The Gompertz distribution has been the workhorse of actuarial science to model mortality since the distribution provides a simple analytic formula for survival based on the observation from many settings that mortality rises exponentially with age (Olshansky and Carnes, 1997). Using simulations, Basu et al. (2004) show that

[53]The estimates are nonetheless similar if I include indicators for individual events rather than the broader event classes.

[54]Medical research links height to an earlier death due to biological factors, such as reduced cell replication and lower cancer incidence (Samaras 2012). The results in Section 5 are not sensitive to including weight, which is highly correlated with height, or body mass index. I include height alone to avoid potential collinearity problems.

the Cox proportional hazard model performs better than log OLS under a Gompertz data generating mechanism, even when there is no censoring. The Cox model is also more efficient in terms of lower root mean square error. While the performance of the Cox model is poor if the proportional hazards assumption is violated, I confirm the proportional hazards assumption is met in my data using tests of Schoenfeld residuals.

The baseline model is the standard Cox proportional hazards model:

$$\lambda = \lambda_0(t)\exp(x'\beta) \tag{2.1}$$

where the hazard of death λ depends on an unspecified baseline hazard $\lambda_0(t)$ and an exponential function of observables. The explanatory variable of interest is an indicator for whether the athlete lost the Olympic final. Robust standard errors are clustered by country and year.

As a parametric alternative to these Cox regression models, I also specify a Gompertz distribution for the baseline hazard. In these models, I allow for individual-level frailty that has a Gamma distribution by specifying the hazard as

$$\lambda = \nu_i \exp(x'\beta + \gamma t) \tag{2.2}$$

I also run models that allow for shared frailty by country. These models assume the unobserved heterogeneity is constant over time. Finally, a robustness test described in Section 2.7 estimates the increasingly mixed proportional hazards model of Frijters et al. (2011) that models unobserved individual heterogeneity as a random walk. The estimates are robust to estimating these various parametric and semi-parametric survival models that make different assumptions about unobserved heterogeneity.

2.5. Results

I first provide non-parametric, unconditional estimates of lifespan that preview my main results. Figure 7 displays Kaplan-Meier survival curves that show Gold medalists die younger than Silver medalists. By age 80, roughly half of Silver medalists remain alive compared to a third of Gold medalists. These differences are statistically significant based on log-rank tests.

Table 14 presents the regression results of the survival models of lifespan described above. The first two columns present Cox proportional hazard models that include event and year fixed effects. Columns 3 and 4 present Gompertz regressions with shared frailty by country. Gold medalists represent the omitted finishing place. In all cases, coefficient estimates are exponentiated and so should be interpreted as hazard ratios. The hazard estimate of 0.570 indicates slightly over half as many Silver medalists are expected to die at any point compared to Gold medalists. The coefficient estimates on losing are statistically significant at the 1 percent level across models. By contrast, estimates on year of birth and ability are not statistically significant.

The results are also robust to including country fixed effects and height as shown in Table 15. This specification relies on within-country variation in winning and losing rather than within-event-year pairs. Columns 1 and 2 present Cox regressions and Columns 3 to 6 present Gompertz regressions that allow for shared frailty by country and event. Across models, the coefficient estimates on losing range between 0.616 and 0.668, again indicating lower hazards of death among losers versus winners.

There is some evidence the relationship between losing and mortality is stronger in specifications that focus on sub-samples of “premiere” events. Although determining which events the public cares most about is somewhat arbitrary, the marquee events in Track are arguably the 100m, 200m, 400m, 1500m, and marathon and the marquee Field events are the decathlon, high jump, and long jump. Table 16 presents Cox regressions based on these sub-samples. The estimates based on within-event-year comparisons (columns 1 and 2) are nearly equivalent to those from Table 14, while the estimates based on within-country identification estimate a stronger relationship between losing and mortality compared to Table 15. While the latter might be interpreted as suggestive evidence that shocks to status are greater in high-profile contests, these regressions largely indicate similar patterns across events.

Similarly, one would expect the association between winning and mortality to be stronger in Olympic Games that were more highly publicized if factors related to status drive the results. To investigate this question, I run regressions that split the sample into two halves before and after 1924. Table 17 shows the results are driven by later Olympic Games, which were more widely covered through print media, radio, and television. The 1924 Paris Games were the first to be broadcast on radio and the 1936 Berlin Games were the first to be televised, for example.[55] In the later period, the hazard estimates of 0.33 on losing indicates that one third of Silver medalists are expected to die relative to Gold medalists at any given time. The estimates on losing from the earlier period are still below 1 but not statistically significant.

[55]The Olympics received more news coverage in later years. Performing a search for articles with the word "Olympics" on the New York Times site during the entire year of an Olympic Games reveals the following counts: 1896: 81, 1900: 36, 1904: 201, 1908: 204, 1912: 533, 1920: 323, 1924: 1,170, 1928: 1,190, 1932: 1,490, 1936: 1,450, 1948: 695. It is not clear why the number of articles drops off in 1948, but one possibility is greater coverage on television and radio. There is a similar pattern in coverage using nationwide results from the website newspaperarchive.com.

2.6. Mechanisms

This section investigates potential mechanisms driving the correlation between Olympic finishing place and mortality. I assess whether mortality is explained by (1) how Olympic finish compares to expectations and (2) income earned after Olympic competition. To study the role of expectations, I construct a metric of expected finishing place based on the performances of each athlete prior to the Olympics and test whether the difference between an athlete's Olympic finish and his expected finish is associated with lifespan. I find that out-performing expectations is positively correlated with lifespan. While Gold medalists either met or exceeded expectations, Silver medalists finished farther ahead of their pre-Olympic ranking. This pattern also operates in reverse for "favorites" who lose: Silver medalists who were previously ranked first die earlier than other Silver medalists. Yet performance relative to expectations cannot fully reconcile the positive correlation between losing and a longer life. More empirical support is found for income as the key mechanism between finishing place and health. To analyze the role of income, I test whether lifespan is associated with income as reported on the 1940 U.S. Census. Losers earned more money than winners and after controlling for income, the estimated mortality hazard does not significantly differ between winners and losers. Higher income has a large, positive, and statistically significant effect on lifespan. Including income also explains roughly twice as much variation in lifespan as the baseline regressions without it. Taken together, these tests suggest that real resources are the key channel between status and health within this sample.

2.6.1. Performance Relative to Expectations

Expectations serve as a natural reference point for judging performance in this setting. As McGraw, Mellers, and Tetlock (2005) argue, psychology suggests that an athlete's (ex ante) expectations affect their perception of their actual performance, ex post. In economic models of expectations-based reference-dependent preferences (Koszegi and Rabin, 2006), measuring expectations is generally difficult outside of the laboratory.[56] The ability to observe recorded performances prior to the Olympics provides an opportunity to cleanly measure expectations in the field. An athlete whose win was expected may receive less of a boost to self-esteem than an athlete who did not even expect to make the final and finished second. To the extent this occurs, one reason Gold medalists die earlier than losers may be that the former were already favored to win and the win represented more of a relief than any positive psychic effect.

To examine the role of prior expectations, I collect the top 100 annual performances by each event. For each event and each Olympic Games, I construct a ranking of the top performers in the 18 months prior to the date of the opening ceremonies of that particular Olympics. An 18 month window provides a long enough window to rank all athletes while still capturing performances close in time to the Olympic Games.[57] Data on performances is collected from the site http://trackfield.brinkster.net and is complete for events going back to 1920, with the exception of the racewalk. I rank unique athletes, not performances, so that only the best performance of an athlete counts towards the ranking.[58] The assumption is that an athlete's expected finish is

[56] See for example Abeler et al. 2011; Crawford and Meng 2011; Card and Dahl 2011; Gill and Prowse 2012 for previous work measuring expectations in laboratory and field settings.

[57] I also obtain similar results if I use a 3-year window instead.

[58] In calculating the pre-Olympic rankings for the 100 meter and 1500 meter runs, I also consider times posted in the 100 yard and mile runs, respectively, since the distances are extremely close. I

based on him running his best time in recent years and all other competitors doing the same.[59]

Table 18 presents the percentage of athletes who were ranked in the top 25, top 10, top 5, top 3, and first prior to the Olympics by finishing place. Gold medalists tended to post the best performances prior to the Olympics. The median pre-Olympic ranking of Gold medalists was 2nd compared to 6th for Silver medalists. Due to positive surprise performances, the average ranking among Gold medalists was 7th prior to the Olympics versus 16th for Silver medalists. Seventy percent of Gold medalists were ranked in the top 3 leading up to the Games compared to 40 percent of Silver medalists. These tabulations of prior performances suggest that Gold medalists likely expected to win more often than Silver medalists. Although being ranked higher before the Olympics increases the chances of victory in the Games "when it counts", success is far from predetermined. This pattern supports the argument that conditional on making an Olympic final, chance plays a key role in assigning status.

I construct an empirical measure of relative performance based on the difference between each athlete's finishing place and his pre-Olympic ranking. Figure 8 plots the distribution of relative performance, defined as pre-Olympic ranking minus Olympic finish. Most athletes were ranked within the top 3. Figure 9 presents a scatterplot of lifespan against this metric of relative performance. There is a positive correlation between lifespan and the numbers of places ahead of expectations an athlete finished, driven by athletes previously ranked outside the top 3. This analysis is now implemented in a regression framework.

subtract 18 seconds from mile times to convert to 1500 meter times and multiply 100 yard times by 1.1 to convert to 100 meter times. These conversions are consistent with the scoring metrics of the International Association of Athletics Federations.

[59] I only observe the top 100 performances by event and year, rather than the full history of each athlete's performances. With the full history, another approach would be to construct distributions of expected finish.

Table 19 presents Cox regressions that include measures of performance relative to expectations. Expectations help to explain the earlier death of winners, even conditional on finishing place. Column 1 reports the specification without the expectations variables for reference.[60] Columns 2 and 3 include the difference between pre-Olympic ranking and Olympic finish. Consistent with the scatterplot, the estimated hazard of death is lower the more that an athlete's Olympic finish exceeds his pre-Olympic ranking. To interpret the magnitude of this estimate, a one-standard deviation increase in this difference is associated with a mortality hazard of 0.734. Including the variable for expectations barely alters the estimate on losing, however, which remains statistically significant. Similar patterns are shown in column 3, which present regressions that include an indicator for whether an athlete was ranked outside the top 20 before the Olympics. The hazard estimate of 0.422 on this variable is lower than the estimate on losing, and both are statistically significant.

To study the effect of "underperforming" on health, I construct an indicator for whether the athlete was ranked first before the Olympics and lost. Intuitively, Silver medalists who were considered favorites were likely more disappointed than athletes who were ranked lower and also lost. The final column of Table 19 display the results of this specification. As shown by the hazard estimates above unity on the interaction term, losing athletes who were ranked first die earlier than other losing athletes. This finding is consistent with favorites who fail to win experiencing more stress or disappointment that is harmful to health compared to Silver medalists with lower expectations.

While performance relative to expectations is positively correlated with lifespan, this metric adds limited explanatory power and cannot fully reconcile the empirical pat-

[60] This is a subsample of that presented in Table 14 because rankings are not available for all events in all years.

tern between winning and a shorter life. The third row from the bottom of Table 19 presents the share of explained variation, similar to an R^2 from a linear regression, as developed by Royston (2006).[61] In the baseline model, 26.6 percent of the variation in lifespan is explained by finishing place and other observables. Including the measure of relative performance increases this share to just 29.4 percent, and the coefficient estimate on finishing place remains large and statistically significant.

2.6.2. Income and Occupational Choices

Real resources like income represent a potentially important mechanism between status and health, either through earnings from winning or by influencing future motivation. The institutional features of this setting make income earned as a direct result of competition limited. Amateurism prevailed until the 1980s and these regulations were strictly enforced, as evidenced by Jim Thorpe—the legendary multi-sport athlete—being stripped of his 1912 Olympic Gold medals for earning money to play minor league baseball in 1909 and 1910 (Flatters, 2000). Most athletes held other occupations while training between Olympic Games.[62] An illuminating account of what could be expected financially after the Olympics comes from the autobiography of Mel Sheppard, a Gold medalist in the 1908 Games. Sheppard describes the parting words he and his Track and Field teammates received from President Theodore Roosevelt after returning from the Olympics during a visit to the White House: "I'm going to give you lads the same friendly bit of advice I gave to my Rough Riders. Remember you're heroes for ten days—when that time's up, drop the hero business

[61] As Royston (2006) describes, this statistic is a modification of that proposed by Nagelkerke (1991) based on the likelihood ratio statistic.

[62] For example, Hannes Koheleman—a Gold medalist distance runner—laid bricks in construction (see The New York Times, "Hannes Kolehmainen, Marathon Champion, is Now U.S. Citizen," January 15, 1921) and Charlie Paddock—a Silver medalist sprinter—worked for a newspaper (see Dallas Morning News, "Obituary: Paddock, Charles William." July 23, 1943).

and go to work" (Sheppard, 1924, p52). The Gold medal itself was worth a modest amount in terms of its metallic content.[63]

While financial rewards from competition were limited, athletes may have pursued various occupations after their athletic career ended, and income from these life decisions may be important to health. Occupational choices after the Olympics help shed light on whether losing serves to motivate. To study this channel, I collect data from the 1940 Census for the sub-sample of U.S. athletes competing between 1920 and 1936. The 1940 Census was the first to record income. Specifically, the survey records annual wage income in 1939 as well as whether the respondent received any supplemental income from other sources. In addition, the 1940 Census also records information on occupation, home ownership, labor supply, race, marital status, and education. The individual records of each Census respondent become publicly available 72 years after the survey, enabling me to observe these variables for my sample of U.S. athletes.

The geneology website Ancestry.com provides digitized Census records from each Census from 1850 through 1940, which can be used to identify specific people based on information recorded in the surveys. To retrieve the records for each U.S. athlete, I first searched using the athlete's name, year of birth, and state of birth. I also followed the "Suggested Hints" provided by Ancestry, which link to other Census records as well as other documents like birth, marriage, and death certificates and army registration cards. These hints are created through a machine learning process and through the family trees built by geneological research that link historical records together. In some cases, the names on the original hand-written Census records are

[63]Before 1912, the gold in the winner's medal was worth about $350 adjusting for inflation and the commodity prices of the year it was awarded (The Economist, 2012). After 1912, gold was no longer used and the winner's medal was made mostly of silver and copper, making it worth even less.

imprecise, leading to the digitized records to be misspelled and requiring additional strategies to search for athletes.[64] To locate athletes whose names do not appear on any of the search returns, I conduct a geographical search that starts with recent known street addresses from either 1930 Census records or army registration cards. The army registration cards also include the date of birth, rather than simply the year of birth as recorded in the Census records, which increases the likelihood of a match along with the athlete's name and place of birth. I then work backwards, manually combing through the the list of Census records from a specific geographical location to retrieve the records of athletes whose names have been misspelled. When street addresses are not available, I begin with all males born in the athlete's state of birth during a 1-year window (older and younger) around the athlete's year of birth.

This process retrieves 80 percent of U.S. athletes in the 1940 Census who competed between 1920 and 1936. This high rate is achieved by a detailed inspection for each athlete based not only on searching by name, but also on geographic and demographic information collected from other biographical sources to narrow the search process. Other studies in economic history that merge individual records across surveys by surname, year of birth, and place of birth, tend to have substantially lower match rates because they use much larger samples of Census data (see e.g. Abramitzky et al. 2012, 2014; Bleakley and Ferrie 2016).

The 1940 Census includes variables for wage income earned in 1939 and whether any income was earned from supplemental sources (yes or no). The survey also collects the number of weeks worked in 1939, number of hours worked the prior

[64] For example, Ancestry's digitized records mistakenly list Edward Gourdin as Edward Gonodin, Robert Van Osdel as Robert Van Vadel, Leo Sexton as Leo Septon, and Raymond Barbuti as Raymond Barbutte.

week, whether the person owns their home or rents, and their stated occupation and industry of employment. The digitized records available on Ancestry include the name of the occupation but not the 3-digit occupation code used for classification, which are instead only listed on the hand-written sheets. I collect the 3-digit occupation codes for each athlete from their original Census records. The 1910, 1920, and 1930 Censuses include occupation and industry but not income. I link athletes to these earlier surveys to record the occupations of their parents.

Earnings by occupation are imputed following two approaches. The first approach uses the average earnings by 3-digit occupation code as reported in the 100% 1940 IPUMS Census file. There are 235 different occupation codes in the 1940 Census classification. The IPUMS mean earnings by occupation include both salary and other income, using data from the 1950 Census, and combine earnings for both males and females. The second approach is based on 15 different industry classifications in 1925 as reported in Margo (2006): gas and electricity; farming; manufacturing; mining; construction; railroad; telephone; wholesale and retail trade; finance, insurance, and real estate; domestic services; medical services; public school teachers; nonprofit services; personal services; and government. This approach relies on assigning the alphabetic occupation and industry fields recorded on the Census records to one of these 15 industries, and so is less precise than using the 1940 numeric occupational codes. While the first approach using 3-digit occupation codes is the preferred method since it is more detailed, the results are nevertheless qualitatively similar in both cases.

Data on occupational choices and average annual earnings by occupation is presented in Table 20. For comparison, the first two columns list the percentage of each occupational category and the average earnings for all U.S. males aged 20 and older in the labor force, respectively. These statistics are tabulated from the 100 percent sample of the 1940 Census made available through the Integrated Public Use Microdata Series (IPUMS). The IPUMS sample, which excludes any identifying information but is useful to gauge broader trends, constructs a measure of average earnings by 3-digit occupation code. Table 20 then aggregates these codes up to broader occupational categories following the IPUMS census classification. At the high end of the earnings distribution, professional workers earned the most money, followed by proprietors, managers, and officials. At the bottom of the income distribution, farm laborers and domestic services earned the least.

The corresponding statistics for Gold and Silver medalists are presented in Columns 3 to 6 of Table 20. While both groups entered occupations that earned substantially more than the U.S. average, Silver medalists chose occupations that paid more than occupations chosen by Gold medalists. The large majority of Silver medalists were classified as Professional Workers and entered occupations, including physicians, with particularly high earnings. Gold medalists were more likely to be classified as Proprietors, Managers, and Officials. A common occupation among Gold medalists was athletic coach, which the Census classifies as a Semi-Professional Worker. In my sample, the average earnings of Silver medalists were 16 percent higher than Gold medalists based on differences in occupational choices.

Table 21 reports results of Cox regressions that include variables for income, labor supply, home ownership, and demographics, as collected from the Census for each athlete in the U.S. sub-sample. Without controlling for income, losing is again correlated with a lower mortality hazard as shown in Column 1. This specification also includes year and event class effects, height, year of birth, and indicators for White and married. The coefficient estimate on losing is large in magnitude and statistically significant, similar to the results from the full sample. Including income in Column

2 drives the estimate on losing closer to 1, and it is no longer statistically significant. By contrast, higher income correlates with a lower mortality hazard. The coefficient estimates on both income variables are large and statistically significant. Including income also explains 57.8 percent of the variation in lifespan as reported by the modified R^2 statistic, compared to just 30.8 percent with finishing place and other observables in Column 1. As shown in Column 3, which includes the number of weeks worked and excludes income, labor supply alone explains far less of the variation in lifespan. Column 4 includes both income and labor supply and demonstrates that income is a strong predictor of mortality. In this sub-sample, the association between winning and an earlier death is explained by higher income among Silver medalists.

To test whether income or expectations are a stronger predictor of lifespan, Columns 6 through 10 include both sets of variables. A few athletes who competed in the racewalk or steeplechase are dropped because historical times are unavailable for these events. Column 6 replicates the main results excluding both expectations and income for comparison. In this sub-sample, outperforming expectations is not correlated with a longer lifespan as shown in Column 7. Income is again negatively correlated with mortality as shown in Columns 8 through 10. These results provide suggestive evidence that real resources are the key mechanism between status and health in this sample.[65]

[65] By comparison, other estimates of the role of income on mortality vary widely based on age and the type of data, ranging from zero to roughly twice as large as my estimates (Smith, 1999; Deaton and Paxson, 2001; Deaton, 2003; Cutler et al., 2011)

Income and occupations of parents

While losers earned more than winners after the Olympics, it is possible that an athlete's occupational choices and earnings were influenced by that of his parents. This section tests whether parental earnings differ systematically across Silver and Gold medalists, which would constitute a source of omitted variable bias if living standards in childhood are correlated with mortality. To study each athlete's family history, I collect their parent's occupations recorded in the 1910, 1920, and 1930 Censuses, when the athletes were in childhood.[66] I impute earnings for their parents' occupations based on the IPUMS occupation codes from the 1940 Census. In the few cases in which a parent held multiple occupations over different waves of the Census, I calculate the average of the two earnings estimates. As a second measure of parental earnings, I also assign the industry-specific average earnings in 1925, which is closer to the time period of Census surveys, collected from Margo (2006). In both cases, parental earnings are not statistically different between Gold and Silver medalists based on a simple t-test (with the parents of Gold medalists having slightly higher earnings). Failing to reject the null hypothesis that parental earnings are equal between winners and losers can be interpreted as another test of balance between the two groups, now on childhood economic conditions.

Table 22 presents results from Cox regressions that include parental earnings along with the individual athlete's income and other Census variables. Since hazard models of mortality cannot include individual fixed effects, these specifications use information about parental earnings to control for unobservables at the family level. Columns 1 to 4 use imputed parental earnings based on 1925 industry occupations and Columns

[66] More specifically, I record the occupation of the household head, which was the father in most cases.

5 to 8 use imputed parental earnings based on 1940 occupational codes. Column 1 includes the athlete's income and that of his parents in logs and Column 2 in levels. The coefficient estimates on parental earnings are not statistically significant while those on the athlete's income once again are. The specifications in Columns 3 and 4 includes the difference between athlete's income and parental income. This measure of the within-family difference in income is positively related to the athlete's lifespan and highly significant. The results are similar when parental income is imputed using 1940 occupation codes in Columns 5 to 8. Taken together, the analysis of Census records are consistent with idea that relative rank influences motivation. The data does not allow me to distinguish whether losing motivates or winning de-motivates, however.

2.6.3. Competing explanations

News coverage

This section tests whether news coverage, which may be correlated with both winning and income, explains lifespan. Even though amateurism prevented athletes being directly compensated for their performance, it is possible that athletes received non-monetary rewards, like housing or job opportunities, that could have first-order effects on longevity. To study this mechanism, I collect text-based data on newspaper coverage of each athlete from the website newspaperarchive.com.[67] I focus on U.S. athletes among the Census sub-sample because the site mainly includes U.S. newspapers and to compare the importance of news coverage against income income as a mechanism. For each athlete, I search for stories containing their first and last name, the word

[67] Other research on news coverage has also used data from this source (Gentzkow et al., 2011)

"Olympics", and the year and event they participated in.[68] I record the number of news stories within two decades of the Olympic Games the athlete competed in. The rationale for restricting coverage to this period is that any changes to living standards as a result of Olympic performance are likely to be reflected in coverage closer to the competition. For example, there was very little newspaper coverage of athletes competing in the first few Olympic Games, but much more coverage in the 1960s and later after most were deceased.[69]

Table 23 reports hazard regressions that include newspaper coverage variables estimated on the U.S. sub-sample with Census records. Columns 1 and 2 present Cox regressions with variables for news coverage along with finishing place, with Column 1 including the count of stories and Column 2 including an indicator for over 50 stories to allow for non-linearity. Losing has a strong and statistically significant association with lifespan, while variables for news coverage do not. Columns 3 through 4 include variables for news coverage and income, omitting finishing place (which is positively correlated with coverage). As before, higher income is associated with lower hazards of death, and the coefficient estimates on news coverage are not statistically significant. Columns 5 and 6 include variables for finishing place, news coverage, and income. These regressions do not suggest that media exposure influences mortality. News coverage also provides limited explanatory power in contrast to income, as shown at the bottom of the table.

---

[68]In case the athlete is known primarily by his nickname, I also include searches that replace the athlete's first name with the nickname reported on sportsreference.com. In addition, since the long jump was historically called the "broad jump" during my sample, I search for this term in that event.

[69]The post-1950s coverage of athletes competing in the first modern Olympic Games tends to recount the experience of these early athletes to establish the history of the Games.

Mean reversion

The correlation between finishing place and mortality is unlikely to be driven by regression to the mean. Due to variability in performance, those ranked below the mean before the Olympics are likely to rank closer to the mean in the Olympic final. The winner in the Olympics may have previously been ranked closer to the mean, finding himself the fortunate recipient of good luck on the day that counts. As discussed earlier, such variation in performance is key to the identification strategy. For mean reversion to explain the relationship between losing and lifespan, however, performance would have to correlate contemporaneously with mortality risk. It seems possible that performance could correlate with other measures of current health, such as resting heart rate or VO_2max, but unlikely it would correlate with long-term health outcomes like mortality.

2.7. Robustness

I perform several checks to verify the validity of the results. First, the results are robust to specifying lifespan using the start date of the Olympic Games rather than the athlete's birthdate, as shown in Table 24. One rationale for using the date of first Olympic competition to "start the clock" is that participation in the Olympics represents the timing of the shock to status. This timing issue is less relevant empirically in my setting where athletes are largely the same age at competition than in the context of Nobel laureates or Oscar winners, where variation in ages of winners and losers is greater. In addition, I test that no single year is overly influential in the analysis by dropping each year in turn (Table 25). I also verify that no single country drives the results by dropping each country in turn (results not shown). In both cases, the main estimates are robust.

Comparing Silver to Gold medalists presents the sharpest cut-off between winning and losing, but the results also hold when including other Olympic finalists who also lost (Table 26). Columns 1 to 4 present Cox regressions that compare the longevity of Gold medalists to that of Silver medalists, Bronze medalists, and 4th place finishers. Again, losing is associated with a lower hazard of death, and the estimates are similar to the main results. Columns 5 through 8 compare Gold medalists to all other finalists (up to 8th place). Gold medalists consistently die earlier than all losers.

Finally, I estimate hazard models that allow for individual time-variant unobserved heterogeneity by specifying the individual heterogeneity as a random walk as in the Increasingly Mixed Proportional Hazards (IMPH) model developed by Frijters et al. (2011). I draw 8000 sample paths of the random walk for each individual and then average the estimates over these draws. As described by Frijters et al. (2011), allowing for time-variant unobservables is intuitively appealing when the researcher observes detailed information about the individual during the baseline period but less about him later in life. This situation characterizes my setting, where there is arguably little unobserved heterogeneity in health status at the time of the Olympic final but many lifestyle decisions and shocks after the competition are unobserved. The coefficient estimates are similar to those presented in the main analysis.

2.8. Conclusion

This paper has compared the longevity of Olympic Track and Field athletes to investigate how competition for status influences health. Counterintuitively, Silver medalists live over two years longer than Gold medalists, on average. The institutional features of the Olympic setting—though highly stylized—allow status to be cleanly identified. The sharp cutoff between winning and losing in the Olympic final helps to reduce possible unobserved heterogeneity between athletes. I also demonstrate that winners and losers are balanced in terms of observables like height, ability, and age, and that awarding status involves a large degree of randomness, likely because it is a physical contest held on a single day every four years. Specific features of the setting are also instrumental in isolating different channels between status and health. In particular, there is less concern of reverse causality here than other contests, at least during the period before the rise of performance-enhancing drugs that I study. The prevailing system of amateurism also prevented any compensation to be earned directly from competition.

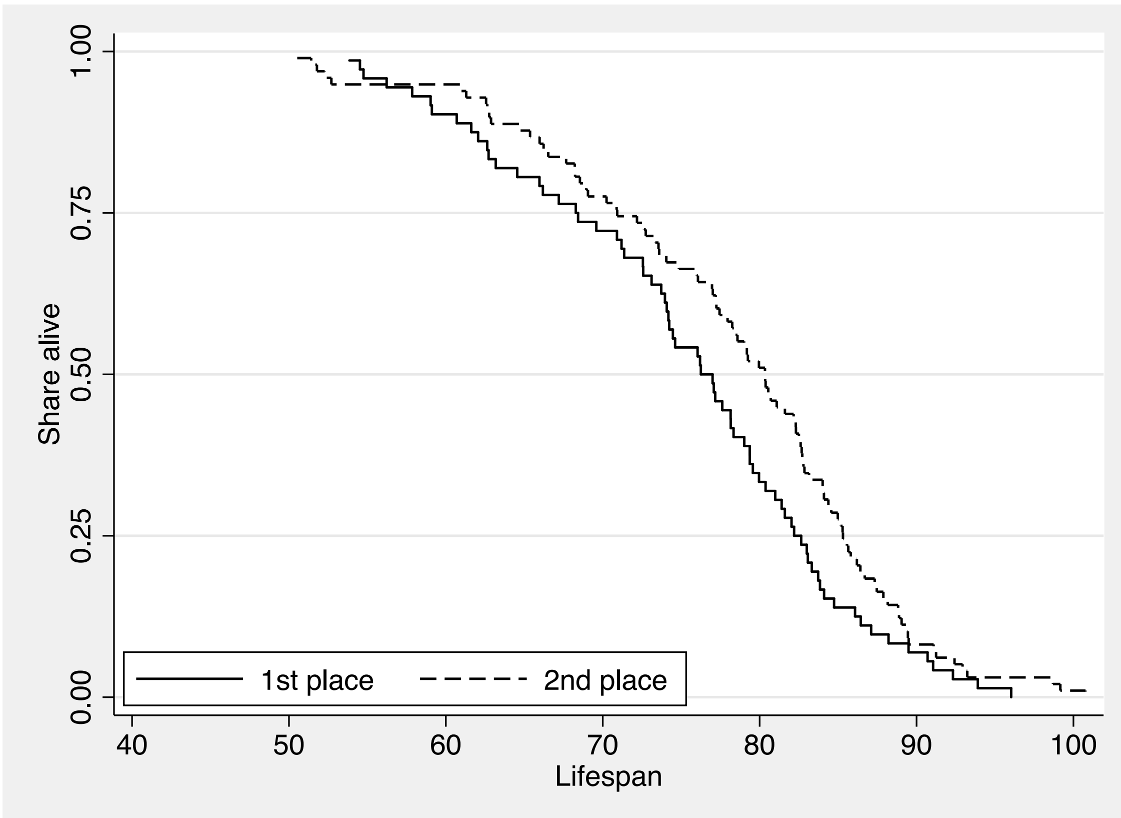
Using individual Census records of each athlete and his family, I find empirical support for occupational choices and income as the key channel between status and health. Silver medalists pursued occupations that paid more money than those chosen by Gold medalists. Income, which is reported in the 1940 Census, is positively correlated with lifespan in the sample and fully accounts for the relationship between losing and mortality. Including income in regression models also explains roughly twice as much variation in mortality as other observables. To test whether childhood economic conditions were similar between Gold and Silver medalists, I link athletes to their family's earlier records in 1910, 1920, and 1930 Censuses and impute parental incomes based on their occupations. The failure to reject the null that parental incomes were equal between Gold and Silver medalists provides another test of balance. Based on a range of specifications and estimation methods, it is unlikely that omitted variable bias explains these patterns. Since amateurism limited any income earned from competition, the analysis of Census records is consistent with relative rank influencing motivation. Data limitations prevent me from determining whether losing motivates or winning de-motivates, however. It is important to note that if Gold medalists enjoyed more income-related opportunities from winning than Silver medalists, such benefits should reduce Gold medalists' mortality risks, not increase them.

There are several limitations of the study. In terms of data, many decisions and shocks after the Olympics are unobserved. The Census records add key information about occupational choices, income, and marital status, but lifestyle factors are likely to matter as well. For example, decisions about smoking and alcohol consumption would be informative to study behavioral responses to winning and losing. I attempt to measure potential psychological responses captured by how performance compares to expectations, and find less support for this mechanism than income. Second, data on each athlete's income after the Olympics is measured only once in 1940. One would ideally observe multiple years of earnings, although lifetime earnings are clearly endogenous. For each athlete's parents, income must also be imputed in the pre-1940 Censuses based on occupation since it is not collected in the survey. Third, the sample size is small, especially on the U.S. sub-sample with Census records. A larger sample could help to improve precision of the estimates and increase confidence in the results. Finally, the setting of Olympic Track and Field raises questions about external validity. This study's findings may be applicable more broadly insofar as the pivotal events in people's lives resemble such competition.

Despite these limitations, this paper's findings challenge conventional wisdom and the conclusions from existing studies that being awarded higher status necessarily improves health (Marmot et al. 1978, 1991; Sylvestre, Huszti, and Hanley 2006; Becker, Chay, and Swaminathan 2007; Rablen and Oswald 2008). Instead, losing can have positive, first-order effects on longevity. The most important trials in our lives often involve a binary outcome, like victory or defeat. This paper's results suggest how people respond to such successes or failures can produce long-lasting consequences for health.

2.9. Tables & Figures

Figure 7: Kaplan-Meier Survival Curves by Finishing Place



Note: This figure plots the proportion of athletes still alive at each age by Gold or Silver medal status. The gap that persists around age 55 between the dotted line (Silver medalists) and the solid line (Gold medalists) illustrates that Silver medalists live longer than Gold medalists.

Table 13: Descriptive Statistics

	1st place (N=72)				2nd place (N=98)				p-value from
Variable	Mean	s.d.	Min	Max	Mean	s.d.	Min	Max	t-test of difference
Lifespan (years)	75.24	10.15	53.85	96.03	78.09	10.56	50.53	100.77	0.078
Distance (1=yes, 0=no)	0.15	0.36	0	1	0.15	0.36	0	1	0.996
Middle-distance (1=yes, 0=no)	0.08	0.28	0	1	0.10	0.30	0	1	0.529
Sprints (1=yes, 0=no)	0.25	0.44	0	1	0.23	0.42	0	1	0.590
Field (1=yes, 0=no)	0.38	0.49	0	1	0.34	0.47	0	1	0.354
Throwing (1=yes, 0=no)	0.11	0.32	0	1	0.14	0.34	0	1	0.423
Year of birth	1898.2	14.94	1869	1926	1899.0	14.34	1871	1925	0.548
Ever set World Record (1=yes, 0=no)	0.03	0.17	0	1	0.04	0.19	0	1	0.642
Height (cm)	181.3	7.78	162	193	180.25	6.69	160	195	0.410
Weight (kg)	75.12	11.94	53	108	74.14	11.78	51	110	0.641

Note: This table displays statistics on lifespan and various observables for Gold and Silver medalists. The final column presents the p-value from the t-test that the means of the corresponding variable are equal between Gold and Silver medalists. Silver medalists live over 2 years longer than Gold medalists and this difference is statistically significant at the 10 percent level. There are not statistically significant differences in the mean year of birth, height, weight, the types of events, or the number of World Record holders between Gold and Silver medalists, demonstrating balance between the two groups of athletes. Height data is available for 133 athletes. Weight data is available for 130 athletes. The other variables are available for 170 athletes.

Table 14: Lifespan Regressions with Event and Year Fixed Effects

	(1) Cox	(2) Cox	(3) Gompertz	(4) Gompertz
Lose (1=yes, 0=no)	0.561***	0.563***	0.551***	0.555***
	(-3.10)	(-3.04)	(-3.03)	(-2.98)
Year of birth	0.959	0.960	0.962	0.963
	(-1.05)	(-1.03)	(-1.16)	(-1.11)
Ever set World Record (1=yes, 0=no)		0.903		0.822
		(-0.30)		(-0.36)
Year effects	Yes	Yes	Yes	Yes
Individual event effects	Yes	Yes	Yes	Yes
Country effects	No	No	No	No
Event class effects	No	No	No	No
Frailty	None	None	Country	Country
Observations	170	170	170	170
Log likelihood	-676.49	-676.47	128.76	128.82

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival model regressions. Losing is defined as finishing in second place. The coefficient estimate below 1 on losing indicates the hazard of death is lower among Silver medalists than Gold medalists. Robust t-statistics clustered by country-year in parentheses, except in shared frailty models. *p<0.1, **p<0.05, ***p<0.01.

Table 15: Lifespan Regressions with Country Fixed Effects

	(1) Cox	(2) Cox	(3) Gompertz	(4) Gompertz	(5) Gompertz	(6) Gompertz
Lose (1=yes, 0=no)	0.668*	0.665*	0.649**	0.655**	0.616**	0.628**
	(-1.93)	(-1.94)	(-2.09)	(-2.01)	(-2.27)	(-2.14)
Year of birth	0.989	0.989	0.988	0.988	0.995	0.995
	(-1.30)	(-1.30)	(-1.25)	(-1.25)	(-0.51)	(-0.48)
Height (cm)	1.015	1.015	1.017	1.017	1.006	1.005
	(0.82)	(0.85)	(1.06)	(1.03)	(0.33)	(0.28)
Ever set World Record (1=yes, 0=no)		1.065		0.904		0.767
		(0.15)		(-0.20)		(-0.48)
Year effects	No	No	No	No	No	No
Event effects	No	No	No	No	No	No
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Event class effects	Yes	Yes	Yes	Yes	Yes	Yes
Frailty	None	None	Country	Country	Event	Event
Observations	133	133	133	133	133	133
Log likelihood	-503.26	-503.25	86.93	86.95	89.05	89.17

Note: This table presents exponentiated coefficient estimates (hazard ratios) from survival model regressions with country fixed effects. All regressions also include indicators for event classes (sprints, middle distance, distance, throws, field, and racewalk), rather than individual events as in Table 2. Similar to Table 2, the coefficient estimate below 1 on losing indicates the hazard of death is lower among Silver medalists than Gold medalists. Robust t-statistics clustered by event-year in parentheses, except in shared frailty models. *p<0.1, **p<0.05, ***p<0.01.

Table 16: Cox Regressions: "Premiere" Events Sub-Sample

	(1)	(2)	(3)	(4)
Lose (1=yes, 0=no)	0.543**	0.557**	0.434**	0.459*
	(-2.10)	(-2.05)	(-1.97)	(-1.71)
Year of birth	1.015	1.025	1.007	1.007
	(0.18)	(0.31)	(0.40)	(0.39)
Ever set World Record (1=yes, 0=no)		0.594		0.591
		(-0.42)		(-0.61)
Height (cm)			0.925**	0.922**
			(-2.02)	(-2.01)
Year effects	Yes	Yes	No	No
Event effects	Yes	Yes	No	No
Country effects	No	No	Yes	Yes
Event class effects	No	No	Yes	Yes
Observations	63	63	53	53
Log likelihood	-183.53	-183.39	-142.89	-142.81

Note: Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions among the sub-sample of events of 100 meters, 200 meters, 400 meters, 1500 meters, marathon, decathlon, high jump, and long jump. The coefficient estimates on losing are slightly smaller to those presented in Tables 2 and 3—which is consistent with these events being considered premiere events in Track and Field—but the differences are not statistically distinguishable. Robust *t*-statistics clustered by country-year (column 1-2) and by event-year (columns 3-4) in parentheses. *$p<0.1$, **$p<0.05$, ***$p<0.01$.

Table 17: Cox Regressions by Year of Olympic Games

	(1)	(2)	(3)	(4)
	Years: 1896-1920		Years: 1924-1948	
Lose (1=yes, 0=no)	0.639	0.645	0.417***	0.408***
	(-1.60)	(-1.60)	(-3.14)	(-3.22)
Year of birth	0.963	0.983	0.860**	0.861**
	(-0.61)	(-0.27)	(-2.02)	(-1.99)
Ever set World Record (1=yes, 0=no)		0.194***		1.398
		(-3.16)		(0.84)
Year effects	Yes	Yes	Yes	Yes
Event effects	Yes	Yes	Yes	Yes
Country effects	No	No	No	No
Event class effects	No	No	No	No
Observations	90	90	80	80
Log likelihood	-298.52	-297.47	-248.53	-248.43

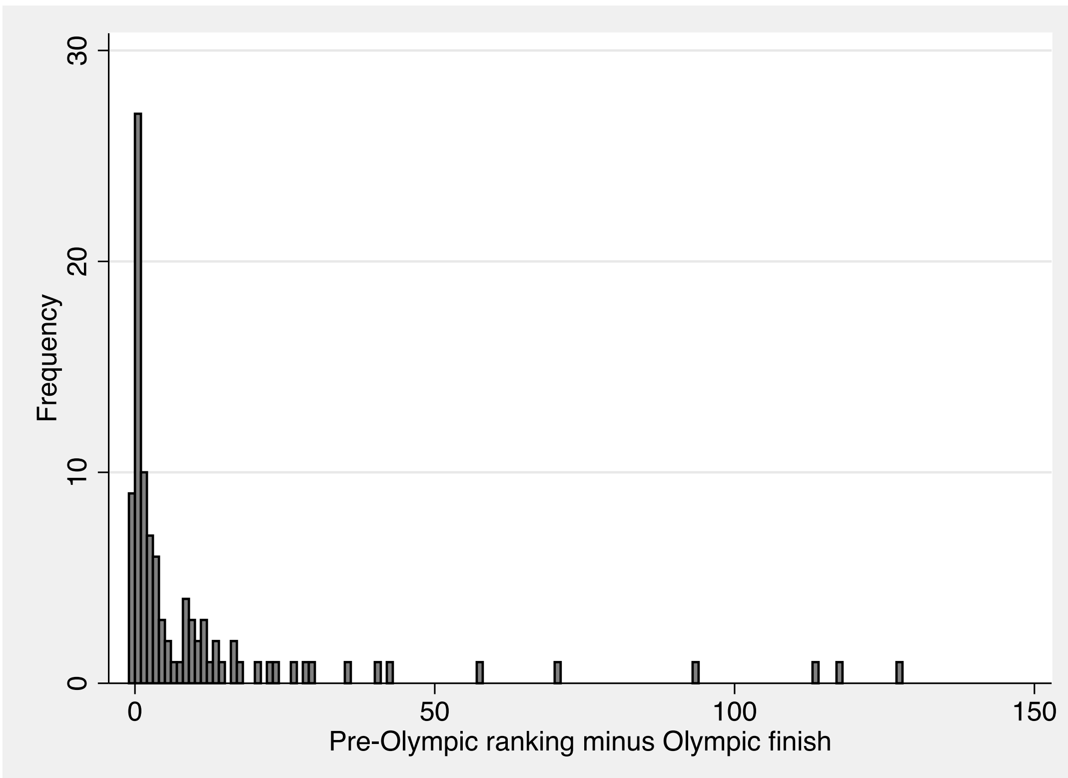
Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions that split the sample by years 1896 to 1920 (columns 1 and 2) and years 1924 to 1948 (columns 3 and 4). Robust t-statistics clustered by country-year in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 18: Statistics of Prior Rankings by Finishing Place

	Gold medalists	Silver medalists
Median Ranking before Olympics	2	6
Mean Ranking before Olympics	7.5	16.4
Top 1 before Olympics (percent)	35.0	15.0
Top 3 before Olympics (percent)	70.0	40.0
Top 5 before Olympics (percent)	77.5	48.3
Top 10 before Olympics (percent)	87.5	58.3
Top 25 before Olympics (percent)	92.5	85.0

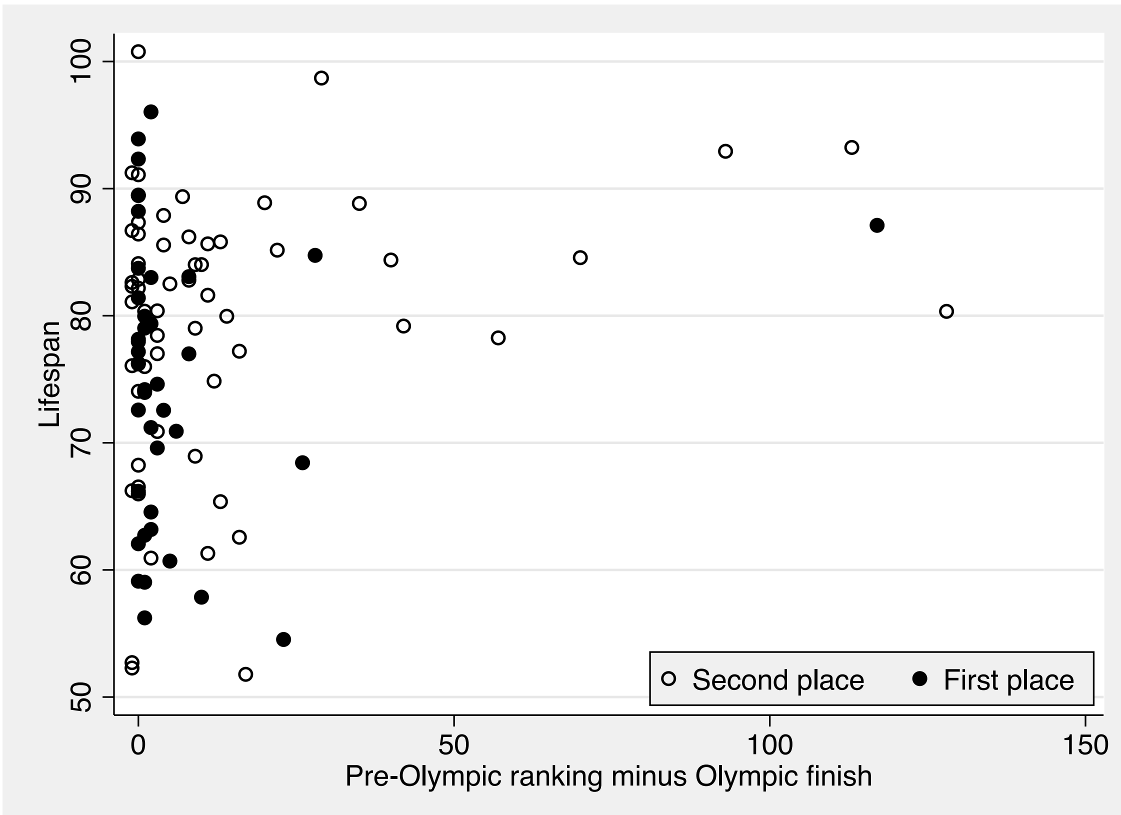
Note: This table displays the average rank and percentage of the sample by their pre-Olympic ranking for Gold and Silver medalists. Rankings are calculated based on the 18 months prior to the opening ceremony of each Olympics.

Figure 8: Distribution of Olympic Finishing Place Relative to pre-Olympic Ranking



Note: This histogram plots the empirical distribution of the difference between pre-Olympic ranking and Olympic finishing place (pre-Olympic ranking minus Olympic finish). Pre-Olympic rankings are calculated using an 18 month window prior to the Olympics. Histogram has unit bin width.

Figure 9: Scatterplot of Lifespan Against Difference Between pre-Olympic Ranking and Olympic Finish



Note: This scatterplot displays the lifespan for each observation in the sample on the vertical axis. The horizontal difference plots the difference between the athlete's ranking based on pre-Olympic performance and their finish in the Olympic games (pre-Olympic ranking minus Olympic finish). Silver medalists are indicated by hollow circles and Gold medalists are indicated by solid circles. The distribution shows that most Gold medalists were ranked within the top-5 before the Olympics, while many Silver medalists were ranked outside the top-5.

Table 19: Cox Regressions: Role of Prior Expectations

	(1)	(2)	(3)	(4)
Lose (1=yes, 0=no)	0.443***	0.455***	0.472***	0.277***
	(-2.78)	(-2.75)	(-2.69)	(-4.30)
Difference between pre-Olympic ranking and Olympic finish		0.989**		
		(-2.43)		
pre-Olympic ranking greater than 20 (1=yes, 0=no)			0.422**	
			(-2.33)	
Ranked 1st before Olympics				0.403**
				(-1.96)
Ranked 1st before Olympics $\times$ Lose				5.055**
				(2.30)
Year of birth	0.964	0.928	0.925	0.977
	(-0.49)	(-1.05)	(-1.08)	(-0.28)
Year effects	Yes	Yes	Yes	Yes
Event effects	Yes	Yes	Yes	Yes
Country effects	No	No	No	No
Event class effects	No	No	No	No
Modified R^2 based on Royston (2006)	0.266	0.294	0.295	0.308
Observations	100	100	100	100
Log Likelihood	-340.32	-337.68	-337.53	-337.03

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions that include variables measuring the athlete's pre-Olympic ranking. The first column replicates the main results from Table 2 on the sub-sample with available ranking data. Column 2 adds the difference between the athlete's pre-Olympic rank and his Olympic finish and Column 3 adds an indicator variable for whether this difference exceeds 20 places. In both Columns 2 and 3, beating expectations is correlated with a lower hazard of death, but the effect of losing remains statistically significant and is of a similar magnitude to the result in Column 1. Column 4 includes an indicator for being ranked 1st prior to the Olympics and the interaction of this variable with losing. The large and statistically significant coefficient estimate on the interaction of 5.055 indicates that "favorites" who lose die earlier than other losers. Robust t-statistics clustered by country-year in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 20: Distribution of Occupations and Average Earnings (1950 US Dollars)

Occupation Category	All U.S. Males aged 20+ in Labor Force		Gold medalists		Silver medalists	
	Percent	Mean Earnings	Percent	Mean Earnings	Percent	Mean Earnings
Professional Workers	4.0	4,340	25.0	4,375	68.0	4,994
Proprietors, Managers, Officials (Except Farm)	8.8	4,090	31.3	4,080	8.0	3,950
Semiprofessional Workers	0.8	3,167	12.5	3,200	4.0	3,200
Craftsmen, Foremen, and Kindred Workers	14.8	3,001	0	-	0	-
Clerical Workers	7.2	2,651	6.3	3,600	8.0	2,500
Salesmen & Saleswomen	5.8	2,547	18.8	2,800	4.0	2,400
Operatives and Kindred Workers	17.0	2,474	0	-	0	-
Protective Service Workers	1.8	2,294	0	-	0	-
Laborers (Except Farm)	13.5	1,968	6.3	2,000	4.0	2,000
Service Workers (Except Domestic and Protective)	4.2	1,696	0	-	0	-
Farmers and Farm Managers	14.1	1,406	0	-	4.0	1,400
Farm Laborers and Foremen	7.5	813	0	-	0	-
Domestic Service Workers	0.5	599	0	-	0	-
Total	100	2,404	100	3,644	100	4,272

Note: This table presents the percent of Gold medalists, Silver medalists, and the male U.S. labor force aged 20 and older by occupational category in 1940 from the U.S. Census. The average earnings by category are constructed from finer 3-digit occupation codes and reported in the IPUMS Census data file. Columns 1 and 2 are tabulated using the 100 percent IPUMS 1940 Census file, which excludes identifying information but provides 3-digit occupation codes and corresponding earnings. Silver medalists pursued higher-paying occupations than Gold medalists.

Table 21: Cox Regressions with 1940 U.S. Census Variables

	U.S. sub-sample: Complete Census data					U.S. sub-sample: Complete Census and pre-Olympics rank data				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lose (1=yes, 0=no)	0.325*	0.543	0.491	0.643	0.591	0.266**	0.270**	0.417	0.392	0.329
	(-1.92)	(-0.98)	(-1.09)	(-0.77)	(-0.88)	(-2.12)	(-2.06)	(-1.36)	(-1.19)	(-1.32)
Log wage income ($100s), annual		0.646**		0.656**				0.645**	0.659*	
		(-2.50)		(-2.34)				(-2.16)	(-1.72)	
Wage income ($100s), annual					0.962***					0.957***
					(-2.68)					(-3.04)
Income from other sources (1=yes, 0=no)		0.067**		0.085**	0.245**			0.055**	0.073**	0.189**
		(-2.57)		(-2.42)	(-2.07)			(-2.57)	(-2.42)	(-2.05)
Owned home (1=yes, 0=no)		1.864	1.312	1.820	1.723			1.632	1.654	1.396
		(0.91)	(0.51)	(0.90)	(0.82)			(0.66)	(0.69)	(0.48)
Number of weeks worked in 1939			1.066	1.061	1.106*				1.082	1.160**
			(1.25)	(1.31)	(1.84)				(1.27)	(2.40)
White (1=yes, 0=no)	0.818	1.401	1.223	1.896	0.847	0.655	0.660	1.562	2.341	0.864
	(-0.39)	(0.76)	(0.22)	(1.10)	(-0.20)	(-0.66)	(-0.64)	(0.71)	(1.42)	(-0.21)
Married (1=yes, 0=no)	0.349*	0.962	0.570	1.472	2.600	0.683	0.611	1.606	6.787	32.684
	(-1.85)	(-0.06)	(-0.59)	(0.57)	(0.92)	(-0.55)	(-0.40)	(0.41)	(1.04)	(1.59)
Difference between pre-Olympic ranking and Olympic finish							1.005	0.970	0.947	0.936
							(0.14)	(-0.92)	(-1.25)	(-1.37)
Height (cm)	1.155***	1.201***	1.144***	1.200***	1.186***	1.160***	1.162**	1.191***	1.181***	1.178***
	(2.79)	(3.40)	(2.75)	(3.32)	(3.40)	(2.69)	(2.46)	(3.34)	(3.30)	(3.46)
Modified R^2 based on Royston (2006)	0.308	0.580	0.351	0.595	0.541	0.284	0.284	0.562	0.595	0.569
Observations	39	39	38	38	38	36	36	36	35	35
Log likelihood	-95.90	-83.52	-90.86	-79.63	-82.47	-86.70	-86.68	-75.30	-70.63	-71.94

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions for the U.S. sub-sample with data from the 1940 Census. All regressions include year and event class effects as well as year of birth (which is not significant). Robust t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 22: Cox Regressions with Parental Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Imputation of parental income based on: 1925 Industry classification				Imputation of parental income based on: 1940 Occupational classification			
		Income measured in:				Income measured in:		
	Logs	Levels	Levels	Levels	Logs	Levels	Levels	Levels
Lose (1=yes, 0=no)	0.520	0.524	0.443	0.456	0.546	0.492	0.468	0.471
	(-1.09)	(-0.80)	(-1.14)	(-1.17)	(-0.99)	(-1.21)	(-0.94)	(-1.00)
Athlete's wage income ($100s), annual	0.597***	0.958*			0.583**	0.952***		
	(-2.63)	(-1.93)			(-2.27)	(-2.97)		
Parent's wage income ($100s), annual	0.861	0.972			0.484	0.943		
	(-0.14)	(-0.28)			(-0.46)	(-1.22)		
Athlete's income minus parent's income ($100s)			0.964***	0.966***			0.970**	0.974*
			(-3.50)	(-3.00)			(-2.57)	(-1.79)
Athlete's income from other sources (1=yes, 0=no)	0.085**	0.277*		0.332	0.063*	0.134*		0.447
	(-2.17)	(-1.68)		(-1.57)	(-1.79)	(-1.78)		(-1.16)
Number of weeks worked in 1939	1.054	1.105	1.105	1.089	1.068	1.155*	1.075	1.064
	(1.10)	(1.54)	(1.60)	(1.36)	(1.27)	(1.73)	(1.21)	(1.04)
White (1=yes, 0=no)	1.650	0.722	0.553	0.626	2.498	1.905	0.399	0.465
	(0.65)	(-0.33)	(-0.53)	(-0.44)	(0.73)	(0.47)	(-0.78)	(-0.60)
Height (cm)	1.245***	1.212***	1.217***	1.199***	1.239***	1.222***	1.217***	1.196***
	(3.64)	(3.41)	(3.11)	(3.47)	(3.97)	(3.92)	(3.13)	(2.99)
Modified R^2 based on Royston (2006)	0.642	0.557	0.509	0.547	0.645	0.582	0.499	0.520
Observations	37	37	37	37	37	37	37	37
Log likelihood	-73.94	-78.59	-80.92	-79.02	-73.71	-77.31	-81.39	-80.37

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions for the U.S. sub-sample with data from the 1940 Census along with parental income based on occupations collected in the 1910, 1920, and 1930 U.S. Census. All regressions include year and event class effects, as well as year of birth, and indicators for married and home ownership (none of which are significant as in Table IX). Parental income does not enter significantly while the variables for athlete income is large and statistically significant. Robust t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 23: Cox Regressions with News Coverage Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	U.S. sub-sample: Complete Census data					
Lose (1=yes, 0=no)	0.271**	0.304**			0.687	0.590
	(-2.13)	(-2.16)			(-0.62)	(-0.93)
Count of news stories	0.997		0.998		0.997	
	(-1.14)		(-0.66)		(-0.78)	
Count of news stories > 50 (1=yes, 0=no)		0.431		1.607		1.261
		(-1.01)		(0.57)		(0.25)
Log wage income ($100s), annual			0.655**	0.654**	0.657**	0.664**
			(-2.49)	(-2.47)	(-2.36)	(-2.26)
Income from other sources (1=yes, 0=no)			0.078**	0.067**	0.084**	0.071**
			(-2.55)	(-2.43)	(-2.44)	(-2.36)
Owned home (1=yes, 0=no)			1.814	1.554	1.856	1.610
			(0.92)	(0.66)	(0.91)	(0.66)
Number of weeks worked in 1939			1.079	1.069	1.067	1.059
			(1.34)	(1.54)	(1.12)	(1.30)
White (1=yes, 0=no)	0.681	0.914	2.258	2.320	2.018	2.130
	(-0.63)	(-0.15)	(1.40)	(1.36)	(1.10)	(1.24)
Married (1=yes, 0=no)	0.344*	0.289*	1.649	1.305	1.542	1.160
	(-1.87)	(-1.92)	(0.76)	(0.38)	(0.63)	(0.20)
Year of birth	0.926	0.938	1.056	1.065	1.039	1.042
	(-0.71)	(-0.50)	(0.41)	(0.51)	(0.28)	(0.32)
Height (cm)	1.168***	1.155***	1.183***	1.203***	1.194***	1.210***
	(3.03)	(2.98)	(3.10)	(2.87)	(3.12)	(3.05)
Modified R^2 based on Royston (2006)	0.348	0.361	0.591	0.593	0.594	0.603
Observations	39	39	38	38	38	38
Log likelihood	-95.27	-95.24	-79.83	-79.65	-79.59	-79.10

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions for the U.S. sub-sample with data from the 1940 Census, along with variables for news coverage. All regressions include year and event class effects. The variables for news coverage do not enter significantly and the estimates for income remain large and statistically significant. Robust t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 24: Robustness Test: Lifespan Calculated as Years post-Olympics

	(1)	(2)	(3)	(4)
	Cox	Cox	Gompertz	Gompertz
Lose (1=yes, 0=no)	0.571***	0.571***	0.543***	0.547***
	(-2.91)	(-2.87)	(-3.28)	(-3.19)
Year of birth	0.859***	0.859***	0.846***	0.847***
	(-3.52)	(-3.52)	(-3.88)	(-3.85)
Ever set World Record		0.991		0.841
(1=yes, 0=no)		(-0.03)		(-0.51)
Year effects	Yes	Yes	Yes	Yes
Event effects	Yes	Yes	Yes	Yes
Country effects	No	No	No	No
Event class effects	No	No	No	No
Frailty	None	None	Individual	Individual
Observations	170	170	170	170
Log likelihood	-671.80	-671.80	60.08	60.13

Note: In these regressions, the dependent variable is years of life calculated beginning from the first Olympic games the athlete competed in, rather than their date of birth. The exponentiated coefficient estimates (hazard ratios) are similar to the main results presented in Table II. Robust t-statistics clustered by country-year in parentheses, except in shared frailty models in columns 3 and 4. *p<0.1, **p<0.05, ***p<0.01

Table 25: Robustness Test: Cox Regressions Excluding Each Year in Turn

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Drop 1896	Drop 1900	Drop 1904	Drop 1908	Drop 1912	Drop 1920	Drop 1924	Drop 1928	Drop 1932	Drop 1936	Drop 1948
Lose (1=yes, 0=no)	0.563***	0.565***	0.504***	0.496***	0.553***	0.576***	0.616**	0.596**	0.572***	0.527***	0.608**
	(-3.01)	(-2.91)	(-3.34)	(-3.81)	(-2.71)	(-2.84)	(-2.56)	(-2.40)	(-2.76)	(-3.32)	(-2.27)
Year of birth	0.978	0.960	0.932	0.961	0.983	0.924**	0.952	0.948	0.967	0.967	0.990
	(-0.56)	(-1.03)	(-1.34)	(-0.99)	(-0.39)	(-2.19)	(-1.15)	(-1.36)	(-0.87)	(-0.72)	(-0.20)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country effects	No	No	No	No	No	No	No	No	No	No	No
Observations	161	164	159	159	151	152	154	150	152	150	148
Log likelihood	-631.35	-649.23	-619.12	-621.86	-581.05	-581.21	-596.54	-578.60	-584.90	-575.90	-570.67

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions that exclude one year at a time from the sample. The results are not sensitive to dropping any single year from the analysis. Robust *t*-statistics clustered by country-year in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 26: Robustness Test: Cox Regressions Including Other Olympic Finalists

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: 1st vs. 2nd - 4th places				Sample: 1st vs. 2nd - 8th places			
Lose (1=yes, 0=no)	0.659***	0.660***	0.695**	0.696**	0.710***	0.710***	0.694***	0.694***
	(-2.99)	(-2.99)	(-2.49)	(-2.49)	(-3.27)	(-3.27)	(-3.30)	(-3.30)
Year of birth	1.004	1.004	0.987***	0.987***	1.010	1.009	0.988***	0.988***
	(0.18)	(0.21)	(-2.96)	(-2.94)	(0.76)	(0.75)	(-3.61)	(-3.60)
Ever set World Record		0.875		0.873		1.051		0.980
(1=yes, 0=no)		(-0.27)		(-0.28)		(0.09)		(-0.04)
Year effects	Yes	Yes	No	No	Yes	Yes	No	No
Individual event effects	Yes	Yes	No	No	Yes	Yes	No	No
Country effects	No	No	Yes	Yes	No	No	Yes	Yes
Event class effects	No	No	Yes	Yes	No	No	Yes	Yes
Observations	350	350	350	350	562	562	562	562
Log likelihood	-1678.24	-1678.19	-1677.79	-1677.73	-2965.38	-2965.37	-2969.43	-2969.43

Note: This table presents exponentiated coefficient estimates (hazard ratios) from Cox regressions that include other finalists. Columns 1 to 4 present results that compare Gold medalists to places 2 through 4 and Columns 5 through 8 compare Gold medalists to places 2 through 8. The results are similar to the main results presented in Tables 2 and 3. Robust *t*-statistics clustered by country-year (Columns 1, 2, 5, 6) or by event-year (Columns 3, 4, 7, 8) in parentheses. *p<0.1, **p<0.05, ***p<0.01.

BIBLIOGRAPHY

Aviva Aaron-Dine, Liran Einav, and Amy Finkelstein. The RAND Health Insurance Experiment, Three Decades Later. *Journal of Economic Perspectives*, 27(1):197–222, 2013.

Aviva Aaron-Dine, Liran Einav, Amy Finkelstein, and Mark Cullen. Moral Hazard in Health Insurance: Do Dynamic Incentives Matter? *Review of Economics and Statistics*, 97(4):725–41, 2015.

Jason Abaluck and Jonathan Gruber. Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. *American Economic Review*, 101 (4):1180–1210, 2011.

Jason Abaluck, Jonathan Gruber, and Ashley Swanson. Prescription Drug Use Under Medicare Part D: A Linear Model of Nonlinear Budget Sets. 2015.

Johannes Abeler and Felix Marklein. Fungibility, Labels, and Consumption. 2013.

Johannes Abeler, Armin Falk, Lorenz Goette, and David Huffamn. Reference Points and Effort Provision. *American Economic Review*, 101:470–492, 2011.

Ran Abramitzky, Leah Platt Boustan, and Katherine Eriksson. Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, 102(5):1832–1856, 2012.

Ran Abramitzky, Leah Platt Boustan, and Katherine Eriksson. A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration. *Journal of Political Economy*, 122(3):467–506, 2014.

Sumit Agarwal, Paige Marta Skiba, and Jeremy Tobacman. Payday Loans and Credit Cards: New Liquidity and Credit Scoring Puzzles. *American Economic Review Papers & Proceedings*, 99(2):412–17, 2009.

Kenneth Arrow. Uncertainty and the Welfare Economics of Medical Care. *American Economic Review*, 53(5):941–973, 1963.

Katherine Baicker, Sendhil Mullainathan, and Joshua Schwartzstein. Behavioral Hazard in Health Insurance. *Quarterly Journal of Economics*, forthcoming, 2015.

Iwan Barankay. Rank Incentives: Evidence from a Randomized Workplace Experiment. 2012a.

Iwan Barankay. Rankings and Social Tournaments: Evidence from a Crowd-Sourcing Experiment. 2012b.

Levon Barseghyan, Francesca Molinari, Ted O'Donoghue, and Joshua Teitelbaum. The Nature of Risk Preferences: Evidence from Insurance Choices. *American Economic Review*, 103(6):2499–2529, 2013.

Anirban Basu, Willard Manning, and John Mullahy. Comparing Alternative Models: Log vs. Cox Proportional Hazard? *Health Economics*, 13:749–765, 2004.

David Becker, Kenneth Chay, and Shailender Swaminathan. Mortality and the Baseball Hall of Fame: An Investigation into the Role of Status in Life Expectancy. *iHEA World Conference*, 2007.

Daniel Benjamin. Do 401(k)s Increase Saving? Evidence from Propensity Score Subclassification. *Journal of Public Economics*, 87(5-6):1259–1290, 2003.

Douglas Bernheim. Taxation and Saving. In Alan Auerbach and Martin Feldstein, editors, *Handbook of Public Economics*, volume 3, pages 1173–1249. 2002.

John Beshears, James Choi, Joshua Hurwitz, David Laibson, and Brigitte Madrian. Liquidity in Retirement Saving Systems: An International Comparison. *NBER Working Paper 21168*, 2015.

Saurabh Bhargava, George Loewenstein, and Justin Sydnor. Do Individuals Make Sensible Health Insurance Decisions? Evidence from a Menu with Dominated Options. *NBER Working Paper 21160*, 2015.

Hoyt Bleakley and Joe Ferrie. Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations. *Quarterly Journal of Economics*, 2016.

Bijan Borah, Marguerite Burns, and Nilay Shah. Assessing the Impact of High Deductible Plans on Health-Care Utilization and Cost: A Changes-in-Changes Approach". *Health Economics*, 20(9):1025–1042, 2011.

Zarek Brot-Goldberg, Amitabh Chandra, Benjamin Handel, and Jonathan Kolstad. What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics, 2015.

Melinda Buntin, Amelia Haviland, Roland McDevitt, and Neeraj Sood. Health Care Spending and Preventive Care in High-Deductible and Consumer-Directed Health Plans. *American Journal of Managed Care*, 17(3):222–230, 2011.

George W. Bush. State of the Union Address, January 2006.

David Card and Gordon Dahl. Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior. *Quarterly Journal of Economics*, 126(1):103–143, 2011.

James Cardon. Health Insurance Choice with Flexible Spending Accounts. *The Geneva Risk and Insurance Review*, 37(2):208–222, 2012.

James Cardon and Igal Hendel. Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey. *RAND Journal of Economics*, 32(3): 408–427, 2001.

James Cardon and Mark Showalter. An Examination of Flexible Spending Accounts. *Journal of Health Economics*, 20(6):935–954, 2001.

James Cardon and Mark Showalter. Insurance Choice and Tax-Preferred Health Savings Accounts. *Journal of Health Economics*, 26(2):373–399, 2007.

Caroline Carlin and Robert Town. Adverse Selection, Welfare, and Optimal Pricing of Employer Sponsored Health Plans. 2009.

Anne Case and Christina Paxson. The Long Reach of Childhood Health and Circumstance: Evidence from the Whitehall II Study. *Economic Journal*, 121(554): F183–F204, 2011.

Amitabh Chandra and Tom Vogl. Rising up with shoe leather? A comment on Fair Society, Healthy Lives (The Marmot Review). *Social Science & Medicine*, 71: 1227–1230, 2010.

Song Chen, Anthony Lo Sasso, and Aneesh Nandam. Who Funds Their Health Savings Account and Why? *International Journal of Health Care Finance and Economics*, 13 (3-4):219–232, 2013.

Raj Chetty. Moral Hazard Versus Liquidity and Optimal Unemployment Insurance. *Journal of Political Economy*, 116(2):173–234, 2008.

Raj Chetty, John Friedman, Tore Olsen, and Luigi Pistaferri. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *Quarterly Journal of Economics*, 126(2):749–804, 2011.

Raj Chetty, John Friedman, Soren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark. *Quarterly Journal of Economics*, 129(3):1141–1219, 2014.

James Choi, David Laibson, and Brigitte Madrian. $100 Bills on the Sidewalk: Violations of No-Arbitrage in 401(k) Accounts. *Review of Economics and Statistics*, 93(3):748–663, 2011.

Andrew Clark and Andrew Oswald. Satisfaction and Comparison Income. *Journal of Public Economics*, 61(3):359–381, 1996.

Philip Clarke, Simon Walter, Andrew Hayen, William Mallon, Jeroen Heijmans, and David Studdert. Survival of the fittest: retrospective cohort study of the longevity of Olympic medallists in the modern era. *BMJ*, 345:e8308–e8308, December 2012.

John Cogan, R. Glenn Hubbard, and Daniel Kessler. *Healthy, Wealthy, and Wise: 5 Steps to a Better Health Care System*. Hoover Institution Press, 2nd edition, 2011.

Alma Cohen and Liran Einav. Estimating Risk Preferences from Deductible Choice. *American Economic Review*, 97(3):745–788, 2007.

Vincent Crawford and Juanjuan Meng. New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income. *American Economic Review*, 101(5):1912–1932, 2011.

Juliette Cubanksi, Christina Swoope, Anthony Damico, and Tricia Neuman. How Much is Enough? Out-of-Pocket Spending Among Medicare Beneficiaries: A Chartbook. Technical report, Kaiser Family Foundation, 2014.

David Cutler, Adriana Lleras-Muney, and Tom Vogl. Socioeconomic Status and Health: Dimensions and Mechanisms. In Sherry Glied and Smith, editors, *The Oxford Handbook of Health Economics*. Oxford University Press, 2011.

Christine Dalton, Gautam Gowrisankaran, and Robert Town. Myopia and Complex Dynamic Incentives: Evidence from Medicare Part D. *NBER Working Paper 21104*, 2015.

Mitch Daniels. Hoosiers and Health Savings Accounts. *Wall Street Journal*, March 2010.

Angus Deaton. Health, Inequality, and Economic Development. *Journal of Economic Literature*, 41:113–158, 2003.

Angus Deaton and Christina Paxson. Mortality, Education, Income, and Inequality among American Cohorts. In David Wise, editor, *Themes in the Economics of Aging*, pages 129–170. University of Chicago Press, 2001.

Mariacristina DeNardi, Eric French, and John Jones. Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy*, 118(1):39–75, 2010.

Jeffrey Dubin and Daniel McFadden. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, 52(2):345–362, 1984.

Richard Easterlin. Will Raising the Incomes of All Increase the Happiness of All? *Journal of Economic Behavior and Organization*, 27:35–47, 1995.

Economist. All that glisters: How much is an Olympic gold medal worth? *The Economist*, April 2012.

Liran Einav, Amy Finkelstein, and Mark Cullen. Estimating Welfare in Insurance Markets Using Variation in Prices. *Quarterly Journal of Economics*, 125(3):877–921, 2010.

Liran Einav, Amy Finkelstein, Stephen Ryan, Paul Schrimpf, and Mark Cullen. Selection on Moral Hazard in Health Insurance. *American Economic Review*, 103(1):178–219, 2013.

Eric Engen, William Gale, and John Karl Scholz. The Illusory Effects of Saving Incentives on Saving. *Journal of Economic Perspectives*, 10(4):113–138, 1996.

Eric Engen, William Gale, and Cori Uccello. The Adequacy of Household Saving. *Brookings Papers on Economic Activity*, 2:65–187, 1999.

C.R. Epstein and R.J. Epstein. Death in The New York Times: The Price of Fame is a Faster Flame. *QJM*, 106(6):517–521, 2013.

Armin Falk and Marcus Knell. Choosing the Joneses: Endogenous Goals and Reference Standards. *Scandinavian Journal of Economics*, 106(3):417–435, 2004.

Martin Feldstein. The Welfare Loss of Excess Health Insurance. *Journal of Political Economy*, 81(2):251–280, 1973.

Martin Feldstein. Rethinking Social Insurance. *American Economic Review*, 95(1):1–24, 2005.

Martin Feldstein. Balancing the Goals of Health Care Financing and Provision. *Health Affairs*, 25(6):1603–1611, 2006.

Fidelity. Retiree Health Costs Hold Steady. Technical report, 2014.

Amy Finkelstein. *Moral Hazard in Health Insurance*. Columbia University Press, New York, 2014.

Ron Flatters. *Thorpe Preceded Deion, Bo*. ESPN, Sports Century, 2000. URL `http://espn.go.com/sportscentury/features/00016499.html`.

Robert Frank. *Choosing the Right Pond*. Oxford University Press, New York, 1985.

Paul Frijters, John Haisken-DeNew, and Michael Shields. The Increasingly Mixed Proportaional Hazards Model: An Application to Socioeconomic Status, Health Shocks, and Mortality. *Journal of Business and Economic Statistics*, 29(2):271–281, 2011.

Paul Fronstin. HSA Balances, Contributions, Distributions, and Other Vital Statistics-A First Look at Data from the EBRI HSA Database on the 10th Anniversary of the HSA. Technical Report 400, Employee Benefits Research Institute, Washington, DC, 2014.

Paul Fronstin. Health Savings Accounts and Health Reimbursement Arrangements Assets, Account Balances, and Rollovers 2006-2014. Technical Report 409, Employee Benefits Research Institute, Washington, DC, 2015a.

Paul Fronstin. Maximizing Contributions to an HSA: Findings from the EBRI HSA Database. Technical report, Employee Benefits Research Institute, Washington, DC, 2015b.

Paul Fronstin, Dallas Salisbury, and Jack VanDerhei. Savings Needed to Fund Health Insurance and Health Care Expenses in Retirement: Findings from a Simulation Model. *EBRI Issue Brief 317*, 2008.

Matthew Gentzkow, Matthew Shapiro, and Michael Sinkinson. The Effect of Newspaper Entry and Exit on Electoral Politics. *American Economic Review*, 101:2980–3018, 2011.

David Gill and Victoria Prowse. A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *American Economic Review*, 102(1):469–503, 2012.

David Gross and Nicholas Souleles. Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence From Credit Card Data? *Quarterly Journal of Economics*, 117(1):149–185, 2002.

Ben Handel and Jonathan Kolstad. Health Insurance for Humans: Information Frictions, Plan Choice, and Consumer Welfare. *American Economic Review*, forthcoming, 2015.

Benjamin Handel. Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review*, 103(7):2643–82, 2013.

Justine Hastings and Jesse Shapiro. Fungibility and Consumer Choice: Evidence from Commodity Price Shocks. *Quarterly Journal of Economics*, 128(4):1449–1498, 2013.

Justine Hastings, Brigitte Madrian, and William Skimmyhorn. Financial Literacy, Financial Education, and Economic Outcomes. *Annual Review of Economics*, 5: 347–373, 2013.

Amelia Haviland, Matthew Eisenberg, Ateev Mehrotra, Peter Huckfeldt, and Neeraj Sood. Do "Consumer Directed" Health Plans Bend the Cost Curve Over Time? *NBER Working Paper 21031*, 2015.

Ori Heffetz. A Test of Conspicuous Consumption: Visibility and Income Elasticities. *Review of Economics and Statistics*, 93(4):1101–1117, 2011.

Ori Heffetz and Robert Frank. Preferences for Status: Evidence and Economic Implications. In Jess Benhabib, Alberto Bisin, and Matthew Jackson, editors, *Handbook of Social Economics*, volume 1A, pages 69–91. Elsevier, The Netherlands, 2011.

Florian Heiss, Adam Leive, Daniel McFadden, and Joachim Winter. Plan Selection in Medicare Part D: Evidence from Administrative Data. *Journal of Health Economics*, 32 (6):1325–1344, 2013.

Lorens Helmchen, David Brown, Ithai Lurie, and Anthony Lo Sasso. Health Savings Accounts: Growth Concentrated Among High-Income Households and Large Employers. *Health Affairs*, 34(9):1594–1598, 2015.

Bob Herman. High-Deductible Plans Dominate Next Open Enrollment. *Modern Healthcare*, November 2014.

Eric Hurst. The Retirement of a Consumption Puzzle. *NBER Working Paper 13789*, 2008.

Emmett Keeler and John Rolph. The Demand for Episodes of Treatment in Health Insurance Experiment. *Journal of Health Economics*, 7(4):337–367, 1988.

KFF. 2014 Employer Health Benefits Survey. Technical report, Kaiser Family Foundation, 2014.

Jeffrey Kling, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel. Comparison Friction: Experimental Evidence from Medicare Drug Plans. *Quarterly Journal of Economics*, 127(1):199–235, 2012.

Peter Kooreman. The Labeling Effect of a Child Benefit System. *American Economic Review*, 90(3):571–583, 2000.

Botond Koszegi and Matthew Rabin. A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics*, 121(4):1133–1165, 2006.

David Laibson. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2):443–478, 1997.

David Laibson, Andrea Repetto, and Jeremy Tobacman. Self-Control and Saving for Retirement. *Brookings Papers on Economic Activity*, 1:91–196, 1998.

Jeffrey Liebman and Richard Zeckhauser. Simple Humans, Complex Insurance, Subtle Subsidies. *NBER Working Paper 14330*, 2008.

Tracy Lieu, Jeffrey Solomon, James Sabin, Jeffrey Kullgren, Virginia Hinrichsen, and Alison Galbraith. Consumer Awareness and Strategies Among Families with High-Deductible Health Plans. *Journal of General Internal Medicine*, 25(3):249–254, 2010.

George Loewenstein, Joelle Friedman, Barbara McGill, Sarah Ahmad, Suzanne Linck, Stacey Sinkula, John Beshears, James Choi, Jonathan Kolstad, David Laibson, Brigitte Madrian, John List, and Kevin Volpp. Consumers' Misunderstanding of Health Insurance. *Journal of Health Economics*, 32(5):850–862, 2013.

Timothy Lu and Olivia Mitchell. Borrowing from Yourself: The Determinants of 401(k) Loan Patterns. 2010.

Annamaria Lusardi and Olivia Mitchell. The Economic Importance of Financial Literacy. *Journal of Economic Perspectives*, 52(1):5–44, 2014.

Erzo Luttmer. Neighbors as Negatives: Relative Earnings and Well-Being. *Quarterly Journal of Economics*, 120(3):963–1002, 2005.

Willard Manning, Joseph Newhouse, Naihua Duan, Emmett Keeler, Arleen Leibowitz, and M. Susan Marquis. Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *American Economic Review*, 77(3):251–257, 1987.

Robert Margo. Wages and Wage Inequality. In Susan Carter, Sigmund Gartner, Michael Haines, Alan Olmstead, Richard Sutch, and Gavin Wright, editors, *Historical Statistics of the United States*. Cambridge University Press, New York, 2006.

M Marmot and A Feeney. General explanations for social inequalities in health. *IARC Scientific Publications*, (138):207–228, 1997.

M Marmot, M Shipley, E Brunner, and H Hemingway. Relative contribution of early life and adult socioeconomic factors to adult morbidity in the Whitehall II study. *Journal of epidemiology and community health*, 55(5):301–307, May 2001.

M G Marmot, G Rose, M Shipley, and P J Hamilton. Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology and Community Health*, 32(4):244–249, December 1978.

M G Marmot, G D Smith, S Stansfeld, C Patel, F North, J Head, I White, E Brunner, and A Feeney. Health inequalities among British civil servants: the Whitehall II study. *Lancet*, 337(8754):1387–1393, June 1991.

Michael Marmot and Eric Brunner. Cohort Profile: The Whitehall II Study. *International Journal of Epidemiology*, 34:251–256, 2005.

A. Peter McGraw, Barbara Mellers, and Philip Tetlock. Expectations and Emotions of Olympic Athletes. *Journal of Experimental Social Psychology*, 41:438–446, 2005.

Victoria Medvec, Scott Madey, and Thomas Gilovich. When Less is More: Counterfactual Thinking and Satisfaction among Olympic Medalists. *Journal of Personality and Social Psychology*, 69(4):603–10, 1995.

Katherine Milkman and John Beshears. Mental Accounting and Small Windfalls: Evidence from an Online Grocer. *Journal of Economic Behavior and Organization*, 71 (2):384–394, 2009.

Alicia Munnell and Anthony Webb. The Impact of Leakages on 401(K)/IRA Assets. Technical Report 2015-2, Boston College Center for Retirement Research, 2015.

Alicia Munnell, Matthew Rutledge, and Anthony Webb. Are Retirees Falling Short? Reconciling the Conflicting Evidence. Technical Report 2014-16, Boston College Center for Retirement Research, 2014.

Nico J.D. Nagelkerke. A Note on a General Definition of the Coefficient of Determination. *Biometrika*, 78(3):691–692, 1991.

Ted O'Donoghue and Matthew Rabin. Doing It Now or Later. *American Economic Review*, 89(1):103–124, 1999.

Council of Economic Advisors. Economic Report of the President. Technical Report H. Doc. 108-145, Washington, DC, 2004.

S. Jay Olshansky and Bruce Carnes. Ever Since Gompertz. *Demography*, 34(1):1–15, 1997.

Stephen Parente and Roger Feldman. Do HSA Choices Interact with Retirement Savings Decisions? In James Poterba, editor, *Tax Policy and the Economy*, volume 22, pages 81–108. 2008.

Mark Pauly. The Economics of Moral Hazard: Comment. *American Economic Review*, 58 (3):531–537, 1968.

Mark Pauly. A Measure of the Welfare Cost of Health Insurance. *Health Services Research*, 4(4):281–292, 1969.

Mark Pauly and John Goodman. Tax Credits for Health Insurance and Medical Savings Accounts. *Health Affairs*, 14(1):125–139, 1995.

Robert Pear. Many Say High Deductibles Make Their Health Law Insurance All but Useless. *New York Times*, 2015. URL `http://www.nytimes.com/2015/11/15/us/politics/many-say-high-deductibles-make-their-health-law-insurance-all-but-useless.html`.

Tony Perrottet. *The Naked Olympics*. Random House, New York, 2004.

Richard Peter and Petra Steinorth. The Optimal Use of Health Savings Accounts: A Consumer's Perspective. 2012.

Richard Peter, Sebastian Soika, and Petra Steinorth. Health Insurance, Health Savings Accounts, and Healthcare Utilization. *Health Economics*, page forthcoming, 2015.

James Poterba, Steven Venti, and David Wise. How Retirement Saving Programs Increase Saving. *Journal of Economic Perspectives*, 10(4):91–112, 1996.

James Poterba, Steven Venti, and David Wise. The Asset Cost of Poor Health. *NBER Working Paper 16389*, 2010.

Drazen Prelec and George Loewenstein. The Red and the Black: Mental Accounting of Savings and Debt. *Marketing Science*, 17(1):4–28, 1998.

Matthew Rablen and Andrew Oswald. Mortality and immortality: The Nobel Prize as an experiment into the effect of status upon longevity. *Journal of Health Economics*, 27(6): 1462–1471, 2008.

Luis Rayo and Gary Becker. Evolutionary Efficiency and Happiness. *Journal of Political Economy*, 115(2):302–336, 2007a.

Luis Rayo and Gary Becker. Habits, Peers, and Happiness: An Evolutionary Perspective. *American Economic Review Papers & Proceedings*, 97(2):487–491, 2007b.

Mary Reed, Ilana Graetz, Vicki Fung, Joseph Newhouse, and John Hsu. In Consumer-Directed Health Plans, A Majority of Patients were Unaware of Free or Low-Cost Preventive Care. *Health Affairs*, 31(2):2641–2648, 2012.

Patrick Royston. Explained Variation for Survival Models. *Stata Journal*, 6(1):83–96, 2006.

Thomas Samaras. How height is related to our health and longevity: A review. *Nutrition and Health*, 21(4):247–261, 2012.

Robert Sapolsky. The Influence of Social Hierarchy on Primate Health. *Science*, 308: 648–652, 2005.

Hendrik Schmitz and Nicolas Ziebarth. Does Price Framing Affect the Consumer Price Sensitivity of Health Plan Choice. *Journal of Human Resources*, forthcoming, 2015.

John Karl Scholz, Ananth Seshadri, and Surachai Khitatrakun. Are Americans Saving "Optimally" for Retirement? *Journal of Political Economy*, 114(4):607–43, 2006.

Hersh Shefrin and Richard Thaler. The Behavioral Life-Cycle Hypothesis. *Economic Inquiry*, 26(4):609–643, 1988.

Mel Sheppard. Spiked Shoes and Cinder Paths: An Athlete's Story. In *Sport Story Magazine*. 1924.

Jonathan Skinner. Are You Sure You're Saving Enough for Retirement? *Journal of Economic Perspectives*, 21(3):59–80, 2007.

James Smith. Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status. *Journal of Economic Perspectives*, 13(2):145–166, 1999.

Petra Steinorth. Impact of Health Savings Accounts on Precautionary Savings, Demand for Health Insurance, and Prevention Effort. *Journal of Health Economics*, 30(2): 458–465, 2011.

Justin Sydnor. (Over)insuring Modest Risks. *American Economic Journal: Applied Economics*, 2(4):177–99, 2010.

Marie-Pierre Sylvestre, Ella Huszti, and James A Hanley. Do OSCAR winners live longer than less successful peers? A reanalysis of the evidence. *Annals of internal medicine*, 145(5):361–363, 2006.

Richard Thaler. Mental Accounting and Consumer Choice. *Marketing Science*, 4(3): 199–214, 1985.

Richard Thaler. Anomalies: Saving, Fungibility, and Mental Accounts. *Journal of Economic Perspectives*, 4(1):193–205, 1990.

Anh Tran and Richard Zeckhauser. Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics*, 96(9):645–650, 2012.

Jenny Tung, Luis Barreiro, Zachary Johnson, Kasper Hansen, Vasiliki Michopoulos, Donna Toufexis, Katelyn Michelini, Mark Wilson, and Yoav Gilad. Social Environment is Associated with Gene Regulatory Variation in the Rhesus Macaque Immune System. *Proceedings of the National Academy of Sciences*, 109(17):6490–6495, 2012.

Ari Vaananen, Aki Koskinen, Matti Joensuu, Mika Kivimaki, Jussi Vahtera, Anne Kouvonen, and Paavo Jappinen. Lack of Predictability at Work and Risk of Acute Myocardial Infarction: An 18-Year Prospective Study of Industrial Employees. *American Journal of Public Health*, 98(12):2264–2271, 2012.

World Anti-Doping Agency WADA. *A Brief History of Anti-Doping.* http://www.wada-ama.org/en/about-wada/history/a-brief-history-of-anti-doping/, 2010.

Gary Wadler. *Doping in Sport: From Strychnine to Genetic Enhancement, It's a Moving Target.* 1998. URL `law.duke.edu/sportscenter/wadler.pdf`.

Anthony Webb and Natalia Zhivan. What is the Distribution of Lifetime Health Care Costs from Age 65? Technical Report 10-4, Boston College Center for Retirement Research, Chestnut Hill, MA, 2010.

Dale Yamamoto. Health Care Costs-From Birth to Death. Technical Report 2013-1, Society of Actuaries, 2013.

Richard Zeckhauser. Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives. *Journal of Economic Theory*, 2(1):10–26, 1970.